MODULATION OF MICAL ACTIVITY BY THE CALPONIN HOMOLOGY DOMAIN: STRUCTURAL, BIOCHEMICAL & THERMODYNAMIC STUDIES

by

Saif Al Qassim

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
March, 2014

Abstract

MICAL (**M**olecule **I**nteracting with **Ca**s**L**) is a 1048 amino acid protein consisting of a flavin-containing monooxygenase domain (FD) with redox activity, a Calponin homology (CH) domain, a LIM (Lin-11, Isl-1, Mec-3) domain, a proline-rich region, and a C-term region containing coiled-coil ERM α-like domain. In axon guidance, MICAL is a *key* molecule that links the extracellular signal from semaphorins –a class of repulsive guidance cues– to the reorganization of the cytoskeleton. Axon guidance is the process by which growing axons respond to extracellular cues that guide them towards their appropriate targets. Proper axon guidance is *vital* in neural development processes such as neuronal cell-migration, axonal branching, path finding, and fasiculation/defasiculation. Our laboratory has previously determined the crystal structure of MICALs FD ($MICAL_{FD}$) and showed that it uses NADPH as the reductant. Studies showed that $MICAL_{FD}$ and $MICAL_{FD\text{-}CH}$ (a construct containing the FD and CH domains only) can bind and oxidize Met44 on actin filaments, thereby affecting their polymerization dynamics. CH domains are typically found in actin-binding proteins. However, modulation of MICAL activities by its non-redox CH domain is poorly understood. In this thesis work, the modulation of MICAL activity by the CH domain was characterized structurally,

biochemically, and thermodynamically. The crystal structure of $MICAL_{FD-CH}$, determined to 2.9-Å resolution, reveals that the CH domain does not interact with the active site in the FD. Furthermore, the FD and CH domains are connected by a disordered linker, which could be important for accommodating the binding of the FD to actin. Kinetic studies reveal that the presence of the CH domain is important for substrate specificity. The structure of $MICAL_{FD-CH}$ was used to build a possible model of the interaction of MICAL with actin. Finally, molecular dynamics simulations using this model revealed that a direct oxidation of actin's Met44 by MICAL is possible, by an extension of the loop containing Met44 of actin. Taken together, this thesis work provides valuable insight into the means by which the CH domain of MICAL modulates its activity.

Advisor and first reader: Dr. L. Mario Amzel

Second reader: Dr. Mario A. Bianchet

Dedication

This thesis is dedicated to my mother and father, whose affection, love, and sacrifices through my whole life made me into the man I am today.

Acknowledgements

This dissertation would not have been possible without the guidance and support of many people who contributed in different ways throughout my time as a graduate student. The people mentioned here have my sincerest gratitude.

First and foremost, my doctorate would not have been possible without the unconditional support of my advisor, Dr. Mario Amzel. My doctorate would not have been possible without him: He showed great faith and belief in me and went to great lengths to ensure that I had the environment, platform, mentorship, guidance, and support needed to succeed. His support from beginning to end has shown me that I have the capacity to be a scientist and pursue an academic career. Dr. Amzel's immense knowledge, which he always shares when mentoring and advising me, has made me realize the qualities an exceptional professor must possess. For that, I am eternally thankful.

I would like to express my deepest gratitude to Dr. Mario Bianchet. Not only did he teach me everything I needed to know in the lab, he has always made himself available for questions and discussions, scientific or otherwise. Most of the work described in this dissertation would not have been possible without his direct guidance. He is a wealth of knowledge, and always takes the time to explain every detail of every concept that I want to learn. Throughout my time in the laboratory,

Dr. Bianchet's mentoring was more like a father passing on knowledge and guidance to his son than than a formal mentor-mentee training. For that, I am deeply indebted to him. I am also very thankful for the help of Dr. Sandra Gabelli, scientifically and personally, during my time in the lab. She is always willing to offer her guidance and support. Like Dr. Amzel and Dr. Bianchet, she has always been available for questions and advice. In the rare instance they do not have all the answers, they would always point me in the right direction. I am really honored to have had the pleasure of being under their guidance.

In addition, I would like to thank all my colleagues in the Amzel laboratory, past and present, for making the laboratory such an exciting place to work in. In addition to being great scientists, they have been very helpful and supportive as scientists and as friends throughout my time in the laboratory. Former laboratory members Dr. Sabri Bora Erdemli, Dr. Krisna Duong-ly, Dr. Simon Messing, Dr. Agedi Boto, Dr. Srinivas Aripirala, Dr. Ignacia Echeverria, Dr. Tzu-lan Yeh, Dr. Danielle Chaves Moreira, Dr. Santiago Di Lella, Dr. Chizu Shimokawa, and Yuly Sanchez have my sincerest gratitude for all the support received while they were in the laboratory. Current laboratory members Olivier Simon, Dr. Anita Ghosh, Dr. Kasia Rudzka, Andres Hernandez, Dr. Mauricio Urquiza, Dr. Michelle Miller, William Hong, Jesse Yoder, and Yunlong Liu have been a pleasure to work with, and have considered me more as a friend than a colleague. Special thanks to Dr. Mauricio Urquiza and Olivier Simon for their engaging discussions and work on the MICAL project, Andres Hernandez with his help performing the thermal shift assay, and

Eitan Borgnia for his help in some of the MICAL experiments. Thank you to all my labmates for great social times together.

I would like to express my gratitude to my thesis committee members Dr. Peter Devreotes, Dr. Susan Craig, and Dr. Mollie Meffert for their support and guidance during thesis committee meetings. Their insight and input was very valuable in this dissertation work. In addition, I am very thankful to Dr. Carolyn Machamer, director of the BCMB graduate program, for her directorship, kindness, advice, support, and allowing me opportunities to actively participate in service and leadership roles in the Hopkins community during my graduate studies. I would also like to thank other faculty members whose discussions, guidance, and advice has helped me during my graduate studies in one way or another: Dr. Wade Gibson, Dr. Jon Lorsch, Dr. Carol Greider, Dr. Jeremy Nathans, Dr. Rachel Green, Dr. Douglas Robinson, Dr. Rob Jensen, Dr. Cynthia Wolberger, Dr. Herschel Wade, Dr. Jungsan Sohn, Dr. Dan Leahy, Dr. Bertrand Garcia-Moreno, Dr. Juliette Lecomte.

Furthermore, my sincerest gratitude goes to all the professors, post-doctoral researchers, and graduate students of the biophysics departments at Homewood and at the School of Medicine. The departments are very collaborative and supportive, and these people make Hopkins biophysics community a great place to be part of.

A special thanks to Tammy Watson, Kathleen Kolish, Ranice Crosby, Teri Pennington, Rhea Dubs, Leslie Brown, Sharon Root, Margie Policastri, Jessica Rexroad, and Arhonda Gogos for all their administrative help. They have been extremely helpful and supportive for my non-scientific issues and questions.

This acknowledgment would be incomplete if I fail to mention the help, support, and encouragement I received from my family: my mother, father, younger brother, sister, youngest brother, and wife. They have provided me with unconditional love and support during my graduate studies and throughout my whole life, and no words can express the gratitude and appreciation I have for that. Last but not least, I am grateful to my daughter Alyazia, who is 1 year and 3 months old at the time of writing this acknowledgement. She is a joy to us all.

Table of Contents

List of Tables

List of Figures

1. Introduction

Our bodies contain billions of neurons and neural connections formed in a very complex yet intricate and specific way. These connections are formed as a result of guidance molecules interacting with growing neurons, directing them towards or away from a certain region, ultimately towards their target tissue. Neurons are able to 'steer' because upon interaction with attractant or repellent guidance molecules, a signaling cascade occurs that leads to rearrangement of the cytoskeletal proteins (actin, microtubules) responsible for maintaining structure, rigidity, and shape of cells. MICAL is a large multidomain flavoprotein whose enzymatic monooxygenase activity is required for neurons' ability to respond to semaphorins, a class of repellent guidance molecules.

1.1 Axon Guidance in the Nervous System

For the intricate network of neural corrections to be formed correctly, axons are guided during growth to their targets by a combination of specific attractive or repulsive extracellular cues. Neurons sense their external environment for these cues through their highly motile and sensitive growth cones, located at the tip of a growing neuron. The growth cones contain plasma membrane receptors that can interact with extracellular cues, and when such an interaction occurs, a signaling cascade is initiated inside the growth cone that leads to rearrangement of the cytoskeleton. As a result, the neuron is 'guided' in response to the extracellular guidance cues, until it reaches its target [1-4].

There are four major families of guidance molecules that have been identified and had their roles established: Netrins, Slits, Semaphorins, and Ephrins. Of these, Semaphorins are the most well-characterized. Semaphorins have been shown to function as repellent axon guidance molecules, and they can be found either as transmembrane or secreted (soluble) proteins. They interact with growth cones of neurons primarily through Plexin receptor family proteins, with Neuropilins functioning as co-receptors [3].

Upon interaction with a semaphorin, Plexin was shown to recruit MICAL to its cytosolic domain [5]. MICAL activity leads to cytoskeletal rearrangement of actin microfilaments, and as a result the growth cone collapses in the region where semaphorin is encountered [6, 7].

1.1.1 Dynamic Modification of Growth Cone Shape

Once the axon is formed from the cell body of the neuron, it grows along a specific path towards target tissue. The shape of the axon is maintained by microtubules. At the growing end of the axon there is a dynamic actin-rich hand-with-finger-like structure known as the growth cone. The growth cone has many protrusions rich in plasma membrane receptors that can interact with external guidance cues (Figure 1), and allow sensing of the external environment. Growth cones change their shape (e.g. stall, collapse, extend, turn) in response to external guidance cues. Once this occurs, microtubules from the main body of the axon are allowed to extend into the growth cone, in the desired direction of growth, thereby stabilizing and extending the axon closer to the target [8, 9].

1.1.1.1 Actin and its role in growth cone shape and morphology

Actin is a globular 374 amino acid ~ 43 kDa protein that can form a double-helical filament from its monomers. The filaments are dynamic in that they continuously grow at one end (plus end) and shrink at the other end (minus end), governed by the nucleotide state of the bound ATP molecule. Filaments can also associate into parallel/anti-parallel bundles or organize into a mesh type lattice. Actin is amongst the most highly conserved protein due to its critical role in structural support and cell movement and dynamics. There are many proteins that bind monomeric (G-actin) or polymeric (F-actin) to regulate actin fiber dynamics [10].

Growth cones are actin-rich at their tips, and F-actin is required for shape change. They possess finger-like filopodia containing parallel F-actin bundles, and lamellipodia containing a mesh type F-actin filament organization (Figure 1). Signaling initiated by receptors in growth cones interacting with guidance molecules induces growth cone shape changes by affecting the actin dynamics. Filament assembly will occur at the sites where attractive guidance cues are encountered, and disassembly will occur at the sites where repulsive guidance cues are encountered [4].

1.2 Introduction to MICAL proteins

Initially identified in T-cells [11], MICAL proteins have also been identified and their importance in control of cytoskeletal dynamics characterized in a variety of neuronal and non-neuronal cell types [12-15]. They have been implicated in processes such as axonal growth cone repulsion, apoptosis, and bristle development in flies. MICAL proteins are large cytosolic proteins with an N-terminal non-heme monooxygenase that contains an FAD cofactor. The FAD co-factor forms H_2O_2 by reducing molecular oxygen, after oxidizing NADPH [16, 17]. In addition to the FD, MICALs contain additional domains known to be involved in protein-protein interaction. These domains may be important for substrate recruitment and modulation of activity. All MICALs contain a CH domain, typically found in proteins that interact with actin [18], and a LIM domain that is implicated in cytoskeletal signaling [19]. Some MICALs contain a coiled-coil ERM domain.

1.2.1 MICAL Proteins Domain Architecture

MICALs are found in invertebrates such as *Drosophila melanogaster* all the way to mammals such as *Mus musculus* (mouse) and *Homo sapiens* (human). *Drosophila* contains one MICAL protein, whereas mice and humans have three. All MICAL-family proteins contain the N-terminal FD, in addition to CH and LIM domains. Most of them have the C-terminal CC domain [13]. In this thesis work, constructs from the *mouse* MICAL-1 protein were used for enzyme kinetics and structure determination, so I will elaborate in more detail the domain architecture

of MICAL-1 only. All subsequent mentioning of MICAL will be in reference to *mouse* MICAL-1, unless indicated.

The N-terminal FD of MICAL spans the first 484 residues, and is separated from the CH domain (residues 511-611) by a 25-residue linker. This domain arrangement and sequence of the FD and CH domains is highly conserved for all MICAL*s* [13]. The LIM domain lies between residues 666-761. This MICAL also contains a CC domain with which it associates with the cytosolic domain of Plexin upon activation by Semaphorins (Figure 2).

1.2.2 Structure of MICAL domains

To date, Nadella et al. and Siebold et al. (2005) have determined high-resolution structures of the FD of *mouse* MICAL-1 by x-ray crystallography [16, 20]. In addition, a solution structure of an isolated *human* MICAL-1 CH domain was reported, and its unfolding was studied by NMR [21, 22].

1.2.2.1 Flavin-containing Monooxygenase Domain (FD)

MICAL's N-terminal flavoprotein monooxygenase domain contains an FAD cofactor shown to oxidize NADPH, with a ~70-fold preference for the phosphorylated dinucleotide, and reduce molecular oxygen to H_2O_2 [16]. Upon reduction by NADPH, the isoalloxazine ring of the FAD adopts an "in" conformation, similar to ρ-hydroxybenzoate hydroxylase (pHBH) and other monooxygenases [20]. In the "in" conformation the reduced FAD, in particular its N5-C4a locus, is less

solvent accessible. Structurally, the FD is related to pHBH; the noteworthy difference being that $MICAL_{FD}$ has a much more open active site and could accommodate a protein substrate, in contrast to the small-molecule substrates in pHBH. The redox activity of $MICAL_{FD}$, studied *in vitro* using kinetic and spectroscopic methods, showed that it exhibits an activity typical of a monooxygenase [17, 23].

1.2.2.2 Calponin Homology (CH) Domain

CH domains are ~100 residues long and have a highly conserved architecture comprised mainly of alpha helices. They carry out diverse functions in cytoskeleton binding and signaling. Protein families that are known to cross-link actin filaments, including α-actinin and spectrin, interact with actin filaments via two CH domains in tandem. These two tandem CH domains are classified as type-1 and type-2. Although type-3 CH domains are found as a single CH domain in proteins such as calponin and IQGAP, proteins containing a single type-2 CH domain also exist [24]. These proteins include MICALs, smoothelins, and RP/EBs. The solution structure of human MICAL-1 CH domain reveals that it exhibits a similar fold and architecture to that of other type-2 CH domains [21]. Its structure is primarily maintained by hydrophobic interactions between the packed residues in the helices [22].

1.2.3 MICAL Effects on Actin Dynamics in Axon Guidance

Mutant *Drosophilla* neurons with mutations in the MICAL gene show severe axon guidance defects. Furthermore, treating rat dorsal root ganglia (DRG) neurons with EGCG – a specific monoxygenase inhibitor – leaves them unable to respond to semaphorin repulsive cues, highlighting the importance of the redox activity of MICAL in axon guidance [5]. It has been recently reported that MICAL's redox activity of constructs containing only the FD alone or FD plus the CH domain can alter actin polymerization dynamics, by specific oxidation of residue Met44 of F-actin, leading to destabilization and breaking up of actin filaments [6, 7]. Addition of F-actin to either *Drosophila* or *human* $MICAL_{FD\text{-}CH}$ was shown to stimulate its redox activity *in vitro*, as measured by NADPH oxidation [7, 17, 23]. Interestingly, although the FD alone is sufficient to bind and oxidize F-actin *in vitro*, the CH domain is required for MICAL effects on actin and proper motor axon guidance *in vivo* [6].

1.3 Thesis Objectives and Summary

Although it is known that the CH domain is functionally important, its role in the MICAL function is unknown. Current structures of the FD and CH domains of MICAL are of individual isolated MICAL domains, and they cannot provide information on how the FD and CH domains interact. This interaction has the potential of controlling and connecting the MICAL activity to interactions with the cytoskeleton. The aim of this thesis is to investigate the modulation of the redox activity of MICAL by the CH domain, using a construct containing the FD and CH domains of MICAL ($MICAL_{FD\text{-}CH}$). Structural (x-ray crystallography, small-angle x-ray scattering), biochemical (absorbance spectroscopy) and thermodynamic (molecular dynamics simulations) methods were applied. Other MICAL truncated constructs were available in our laboratory: a construct containing the FD only ($MICAL_{FD}$, residues 1-484), a construct containing the FD, CH and LIM domains ($MICAL_{FD\text{-}CH\text{-}LIM}$, residues 1-761), a construct of the C-terminal region of MICAL ($MICAL_{C\text{-}term}$, residues 834-1048), and a construct of full-length MICAL (fl-MICAL).

Kinetic experiments performed to follow the redox activity of $MICAL_{FD}$ and $MICAL_{FD\text{-}CH}$ (chapter 2) suggest that the CH domain does not modify the active site of the enzyme, and that it is involved in substrate specificity.

In chapter 3, the crystal structure of $MICAL_{FD\text{-}CH}$ is reported. Both the FD and CH domains adopt a similar fold to that seen in structures of the isolated domains. The CH domain contacts the FD far from where NADPH is predicted to bind and does not modify the predicted NADPH binding site, in agreement with no variation

of the K_M for NADPH observed in the steady state kinetics results. Furthermore, the two domains are connected by a flexible (disordered) linker, and this may be important for optimizing binding to actin.

In chapter 4, molecular dynamics simulations were performed in order to test whether the loop of actin that contains Met44, the residue oxidized by MICAL, can adopt a conformation where it is in close proximity to the FAD cofactor. Results from the simulations show that Met44 can reach a position within the FD where it can be oxidized by the FAD-reduced O_2.

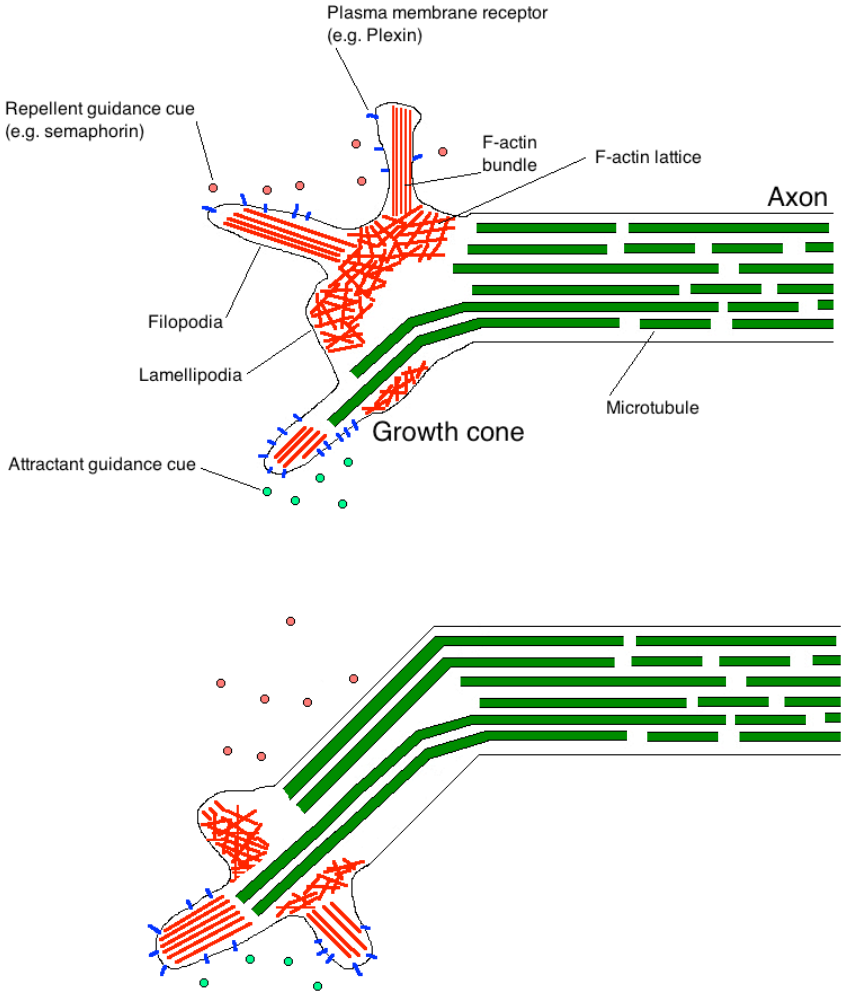


Figure 1. Axon guidance.

Turning of an axon upon growth cone response to attractant and repellent guidance cues (Adapted from [8]).

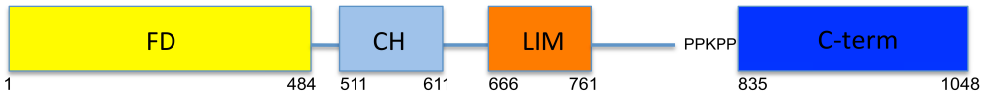


Figure 2. Domain organization of *mouse* MICAL-1.

Numbers indicate amino acid numbers where domains are located in the full-length protein. The protein is 1048 amino acids long.

2. Enzymatic Activity of $MICAL_{FD\text{-}CH}$

In this thesis, I sought to determine the modulation of the redox activity of the FD of MICAL by its CH domain, by comparing the redox activity of $MICAL_{FD\text{-}CH}$ with $MICAL_{FD}$. Kinetic experiments were used to follow the redox activity, and only results for $MICAL_{FD\text{-}CH}$ will be presented in this chapter (M.U. performed the same experiments outlined in this chapter with the $MICAL_{FD}$ fragment). A comparison of the difference in the activities between $MICAL_{FD\text{-}CH}$ and $MICAL_{FD}$ will be discussed in the text.

2.1 Introduction

The FD of MICAL is a monooxygenase domain containing a flavin-adenine dinucleotide (FAD) cofactor, structurally similar to pHBH [16, 20]. The noteworthy difference between MICAL and pHBH is the presence of protein-protein interaction domains in MICAL (Figure 2). The domain immediately following the FD in MICAL is the CH domain. CH domains are found in proteins that bind the actin cytoskeleton [18], but the function of MICAL's CH domain was unclear prior to the work in this thesis. The most straightforward method to determine the effect of the CH domain on the redox activity of the FD is to use a fragment of MICAL containing the FD and CH domains ($MICAL_{FD\text{-}CH}$) and to test *in vitro* the redox activity of this fragment, and compare it with a MICAL fragment containing only the FD ($MICAL_{FD}$).

A simple kinetic assay was used to monitor the redox activity of $MICAL_{FD\text{-}CH}$. The FAD in the FD needs to be reduced by NADPH to activate the enzyme. Thus, in a solution containing $MICAL_{FD\text{-}CH}$, NADPH, and a substrate, NADPH is consumed, with the rate of consumption being directly proportional to the rate of the enzymatic activity. NADPH has an absorbance peak at 340 nm (ε_{340} = 6.2 $mM^{-1}cm^{-1}$), while its product $NADP^+$ does not absorb at that wavelength. It is this attribute of NADPH that allows us to follow the reaction in real-time by ultraviolet (UV) spectroscopy.

2.2 Results

All enzyme kinetics experiments were carried out using either O_2 alone or O_2 and actin as substrates and NADPH as the source of reducing equivalents. The reaction was followed at 340 nm at 298K with a JASCO-v560 UV-Visible Spectrophotometer. Assays were performed in 110 µl of a solution, with the buffer for each experiment mentioned below. To estimate Michaelis-Menten parameters K_M and k_{cat}, NADPH was added at concentrations from 10 to 180 µM, and the initial rates from the first 60 seconds of the reaction were fit to the Michaelis-Menten equation using gnuplot non-linear curve fitting:

$$rate = \frac{V_{\max}[S]}{K_m + [S]} \tag{2.1}$$

where the *rate* is that of NADPH consumption by $MICAL_{FD\text{-}CH}$, V_{max} is the maximal rate of the reaction, and K_M is the substrate concentration [S] at which rate of the reaction is one half of V_{max}. In the experiments outlined below, [S] is [NADPH].

2.2.1 Redox activity

The experiment was carried out in a buffer previously determined in our laboratory to be optimal for following MICAL redox activity: 10 mM Tris-HCl pH 7.0, 20 mM NaCl [16]. 10 µl of 600 nM of $MICAL_{FD\text{-}CH}$ was used in each measurement, with 100 µl of vaying NADPH concentrations. Figure 3 shows a typical dataset, with a k_{cat} of 1.1 ± 0.03 s^{-1} and a K_M of 5.6 ± 1.0 µM. This is on the same order of magnitude to that of $MICAL_{FD}$: k_{cat} of 1.8 ± 0.1 s^{-1}, and a K_M of 3.0 ± 0.7 µM.

These results indicate that the presence of the CH domain does not affect the binding of NADPH to the active site in the FD, as both $MICAL_{FD}$ and $MICAL_{FD\text{-}CH}$ have a similar K_M. Furthermore, the fact that both these fragments share a similar k_{cat} suggests that the active site is unperturbed and has a similar conformation in $MICAL_{FD\text{-}CH}$ compared with $MICAL_{FD}$.

2.2.2 Inhibition by $MICAL_{C\text{-}term}$

Previous studies of MICAL truncation mutants *in vivo* provided indication that the C-term region (immediately following the LIM domain, aa 761-1048) of MICAL can inhibit the redox activity [25, 26]. The intra-molecular control of redox activity is important for two main reasons. First, MICAL can reduce O_2 and produce H_2O_2 using NADPH reducing equivalents [16]. Although H_2O_2 itself is not a strong oxidant, hydroxyl radicals (reactive oxygen species – ROS) can be generated from it upon photolysis, interaction with reactive metals (e.g. Iron), or activation by thiols of cysteine (methionine) side chains. ROS are very reactive and can cause oxidative damage to macromolecules that could lead to harmful effects to the cell even at moderate concentrations (~100 μM) [27]. Second, MICAL directly oxidizes actin filaments via its monooxygenase activity, leading to filament disassembly, upon activation by Semaphorin-Plexin signaling [5, 7]. As a result, the growth cone collapses in the region where the repellent semaphorin signal is encountered. If MICAL cannot be controlled by the signaling pathway, it could lead to axon guidance

defects (as a result of not responding appropriately to the attractant or repellent guidance cues) in those neurons.

The effect of the C-term of MICAL on the redox activity was tested with the steady-state kinetic assay, by addition of a MICAL fragment containing the C-terminal residues 824-1048 ($MICAL_{C\text{-}term}$) to $MICAL_{FD\text{-}CH}$. The same experimental conditions were used as in the previous section, varying the concentration of $MICAL_{C\text{-}term}$.

Results show that addition of $MICAL_{C\text{-}term}$ to $MICAL_{FD\text{-}CH}$ leads to a decrease in the redox activity; the same thing is observed for $MICAL_{FD}$. It was observed that addition of $MICAL_{C\text{-}term}$ decreases the V_{max} but not the K_M; indicative of *non-competitive* inhibition (Figure 3). The data were fit to the modified Michaelis-Menten equation for non-competitive inhibition to estimate the inhibition constant K_i. $MICAL_{FD}$ and $MICAL_{FD\text{-}CH}$ have a similar K_i of inhibition by $MICAL_{C\text{-}term}$: 138 ± 18 μM and 128 ± 8.6 μM for the two fragments, respectively.

2.2.3 Activation by actin

Prior to the experiments in this section being conducted, it was known that actin is a substrate of MICAL and that MICAL can modify actin filaments, leading to their disassembly [7]. However, the kinetic parameters of this oxidation were unknown.

To determine these parameters, the kinetic assay following NADPH consumption was carried out as in 2.2.1, for different concentrations of globular

actin (G-actin, monomer) or filamentous actin (F-actin, polymer). Actin from Rabbit Skeletal Muscle was used (Cytoskeleton Inc., cat. # AKL99), which was initially reconstituted to 100 µl at a concentration of 222 µM.

For G-actin experiments, buffer G was used (5 mM Tris-HCl pH 8.0, 0.2 mM $CaCl_2$). The 222 µM actin was diluted to 22 µM in buffer G, incubated on ice for 30 mins, and ultracentrifuged 100,000 g for 30 minutes at 4˚C to pellet any nuclei; the supernatant was carefully removed for use in the experiments.

The experiments with F-actin were also in buffer G, with the addition of 50 mM KCl, 2 mM $MgCl_2$, 1 mM ATP to maintain polymerization conditions. The 222 µM actin was diluted to 22 µM in the aforementioned buffer, and allowed to polymerize for 30 mins at room temperature, before use.

Addition of G-actin or F-actin to the $MICAL_{FD\text{-}CH}$ and NADPH reaction shows an increase the k_{cat} on the order of 5-fold, compared with no actin (Figure 4, 5). This increase is similar whether G-actin or F-actin was added. This is not the case with $MICAL_{FD}$, where only a small increase in k_{cat} is observed.

This result suggests that the CH domain is responsible for substrate specificity, because only when it is present in MICAL the redox activity is enhanced by actin. In the absence of actin and with NADPH present, the enzyme produces 1 equivalent of H_2O_2 per molecule of NADPH. To test whether the reducing equivalents from NADPH are going directly to actin or towards bulk H_2O_2 production when actin is added, a preliminary experiment was performed where the total [H_2O_2] was measured at different F-actin concentrations using the Amplex Red Hydrogen Peroxide/Peroxidase Assay (A22188, Sigma) (ε_{560} = 44.3 $mM^{-1}cm^{-1}$).

The experiment shows a decrease in total H_2O_2 produced upon adding increasing concentrations of F-actin (Figure 6). To discount the effect of bulk H_2O_2 oxidation of actin non-enzymatically, a control experiment was carried out where H_2O_2 was added instead of NADPH (Figure 6). Taken together, the experiment shows a decrease in H_2O_2 upon addition of F-actin: 2 molecules less per molecule of F-actin to be precise. This suggests that the enhanced activity, and reducing equivalents from NADPH, go directly towards modifying actin.

2.2.4 Effect of NaCl

Ionic strength in the cytosol can vary, so to test the effect of ionic strength on the redox activity of MICAL, NaCl was added to a reaction containing $MICAL_{FD\text{-}CH}$ and varying NADPH concentrations.

The result shows that the major effect is on the K_M for NADPH: K_M increases with increasing NaCl concentration (Figure 7). The same result is observed for $MICAL_{FD}$. This means that NaCl is competing with NADPH for binding, or in other words, NaCl competes for the same binding site or region of MICAL as NADPH.

2.3 Conclusions and Discussion

The described experiments provide important clues about the possible roles of the CH domain. First, given that $MICAL_{FD}$ and $MICAL_{FD\text{-}CH}$ have similar redox activity with respect to NADPH consumption, it suggests that the presence of the CH domain does not change the active site of the FD. Furthermore, since $MICAL_{FD}$ and $MICAL_{FD\text{-}CH}$ have similar values of K_M for NADPH, the conformation of the active site of the FD in $MICAL_{FD}$ and $MICAL_{FD\text{-}CH}$ is the same.

When no actin is present in the reaction with $MICAL_{FD}$ or $MICAL_{FD\text{-}CH}$, their respective k_{cat}s of NADPH consumption are similar. Upon addition of F-actin, only $MICAL_{FD\text{-}CH}$ shows an increase in k_{cat}, but not $MICAL_{FD}$. $MICAL_{FD\text{-}CH}$ binds to F-actin with low μM affinity, as determined by a co-sedimentation approach (Figure 8). Taken together, this strongly suggests that the CH domain plays a role in substrate recognition and specificity, such that the redox activity of the enzyme increases only when the substrate (F-actin or G-actin) is present. At the time of writing this dissertation, a study was reported where authors observed a similar enhancement for *human* $MICAL_{FD\text{-}CH}$ in presence of F-actin [28].

The enhancement of $MICAL_{FD\text{-}CH}$ redox activity observed (with F-actin and/or G-actin) here is not in agreement with other studies [7, 17, 23]. In the presence of F-actin but not G-actin, Zuchinni et al. and Mcdonald et al. report an enhancement in the activity of *human* MICAL-1 and MICAL-2 activity, respectively. They used $MICAL_{FD}$ constructs that lack the CH and the rest of the MICAL domains. This may be due to the requirement of MICAL for the CH domain to be able to recognize G-actin

as a substrate. Hung et al. use a *Drosophilla* $MICAL_{FD\text{-}CH}$ construct and also show that, with F-actin but not G-actin, the NADPH consumption by MICAL is enhanced. A possible explanation is that *Drosophilla* MICAL and *Drosophilla* actin used in that study may behave differently from mammalian (*mouse* MICAL-1) used in this study.

Less bulk H_2O_2 is produced by $MICAL_{FD\text{-}CH}$ in the presence of F-actin, and this suggests that the oxygen atoms of the reduced C4a-hydroperoxidate FAD intermediate are used to oxidize the actin filaments. It remains to be tested whether this modification is a direct oxidation, or as a result of increased local H_2O_2 concentration near the actin filament.

The fact that NaCl has a similar effect on $MICAL_{FD}$ and $MICAL_{FD\text{-}CH}$, namely, that it increases the K_M for NADPH consumption, further supports the conclusion that the FD in $MICAL_{FD}$ and $MICAL_{FD\text{-}CH}$ shares the same conformation.

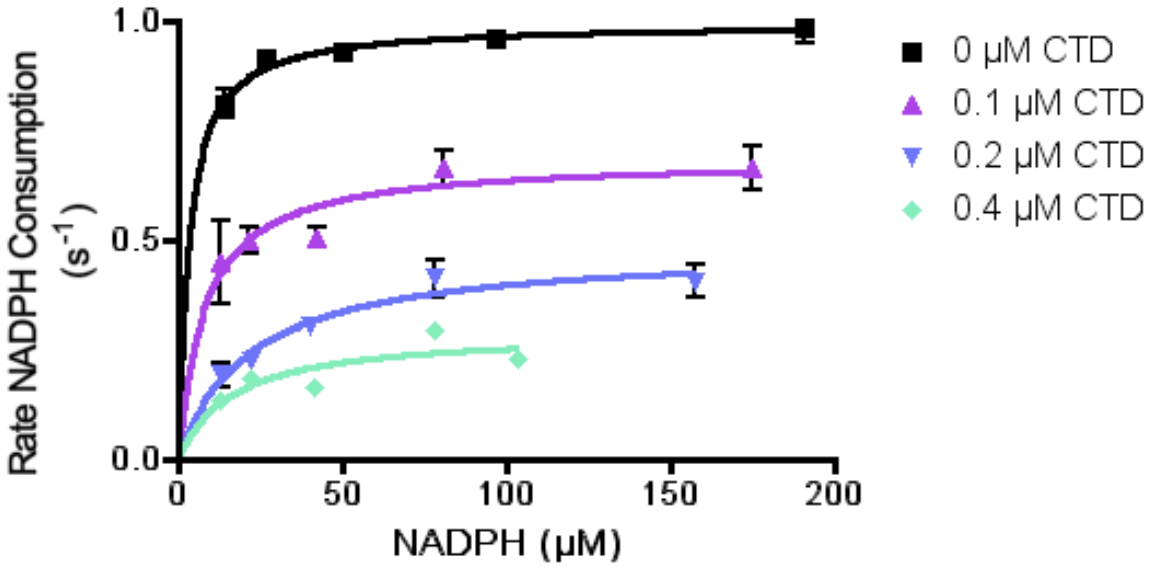
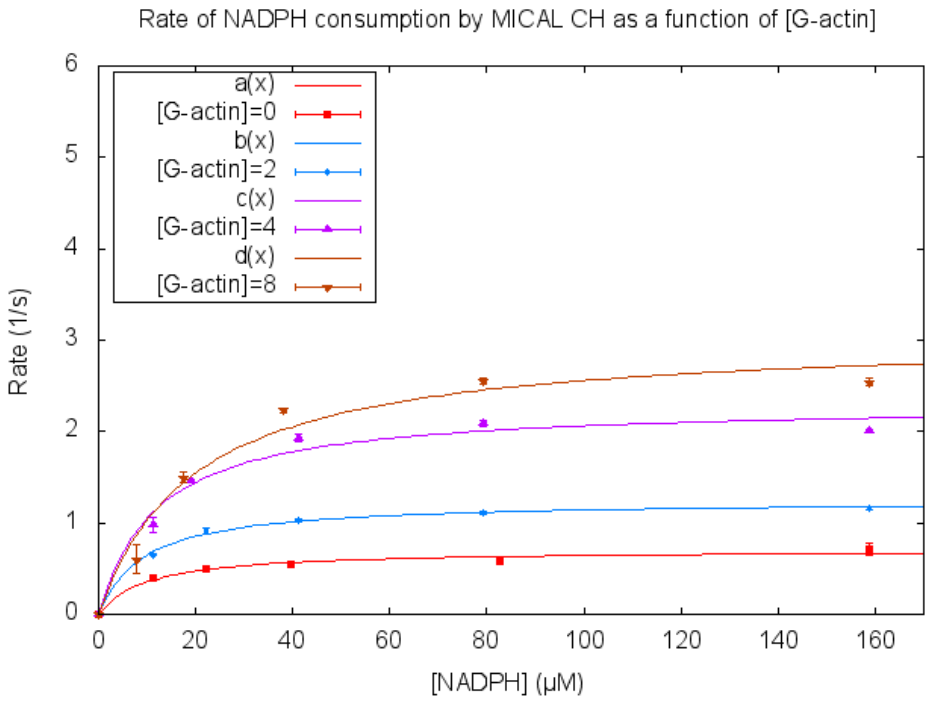


Figure 3. $MICAL_{FD\text{-}CH}$ inhibition by $MICAL_{C\text{-}term}$.

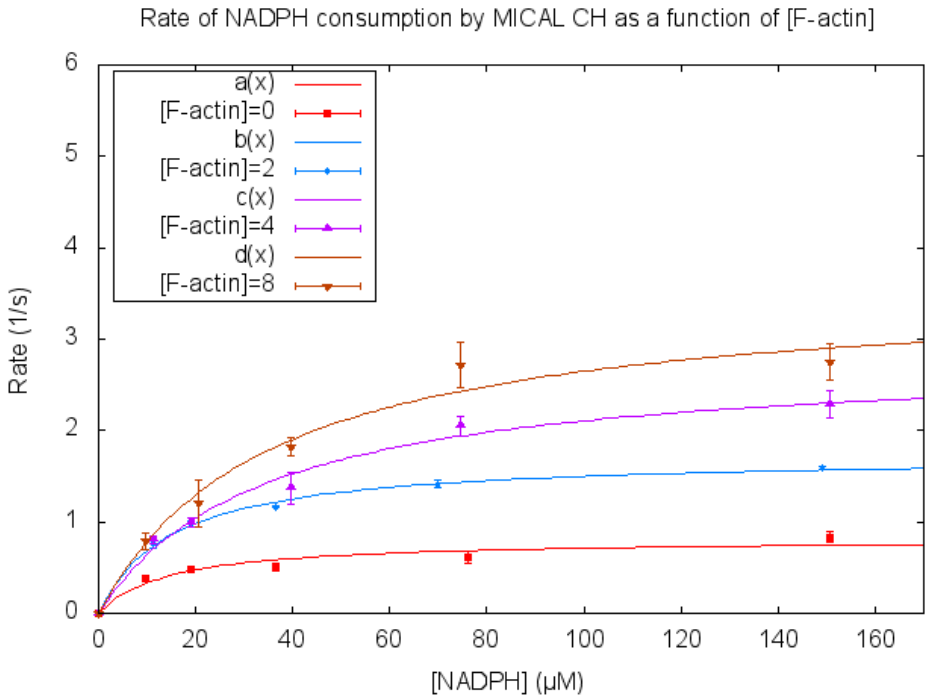
CTD – C-terminal domain of MICAL. Concentrations of $MICAL_{C\text{-}term}$ added are indicated in the legend. The graph was prepared using PRISM software.



[G-actin] (μM)	K_M (μM)	k_{cat} (s^{-1})
0	9.8 ± 2.5	0.7 ± 0.04
2	9.2 ± 0.8	1.2 ± 0.02
4	11.8 ± 2.8	2.3 ± 0.1
8	19.5 ± 5.6	3.1 ± 0.3

Figure 4. $MICAL_{FD\text{-}CH}$ NADPH consumption at different G-actin concentrations.

Rate of NADPH consumption was measured, for different NADPH concentrations, in the presence of 0 μM (red), 2 μM (blue), 4 μM (purple), and 8 μM (brown) of G-actin. Plot prepared using gnuplot software. The table shows the estimated parameters K_M and k_{cat} for each curve ± standard error.

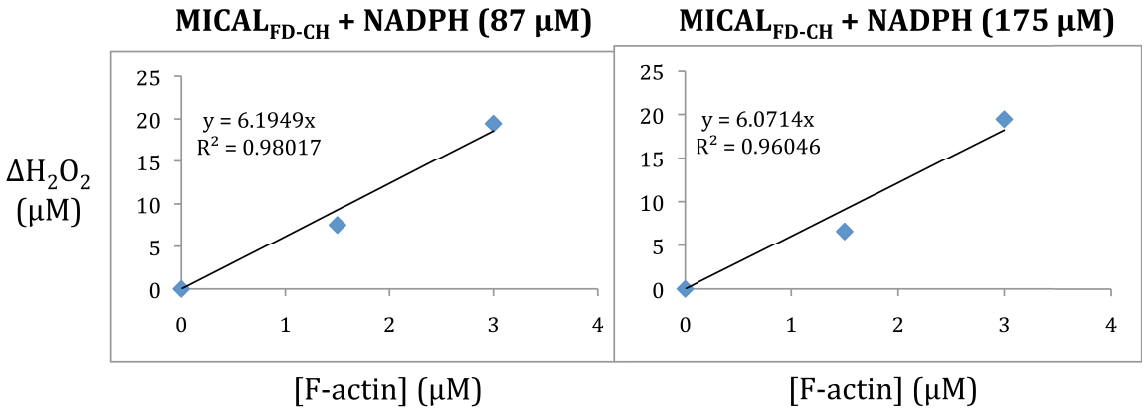


[F-actin] (μM)	K_M (μM)	k_{cat} (s^{-1})
0	14.3 ± 5.7	0.8 ± 0.09
2	15.1 ± 1.3	1.7 ± 0.04
4	33.9 ± 6.3	2.8 ± 0.2
8	35.3 ± 8.8	3.6 ± 0.3

Figure 5. $MICAL_{FD\text{-}CH}$ NADPH consumption at different F-actin concentrations.

Rate of NADPH consumption was measured, for different NADPH concentrations, in the presence of 0 μM (red), 2 μM (blue), 4 μM (purple), and 8 μM (brown) of F-actin. Plot prepared using gnuplot software. The table shows the estimated parameters K_M and k_{cat} for each curve ± standard error.

A



B

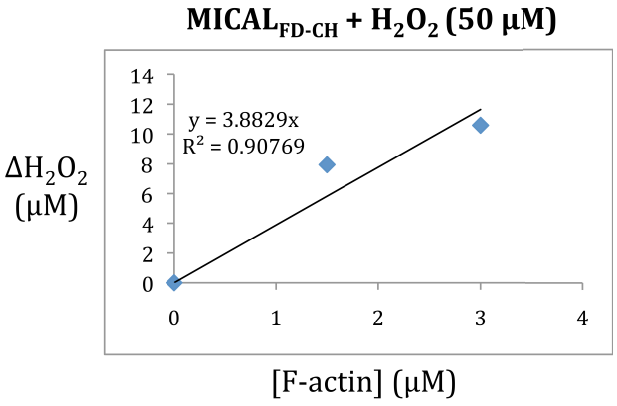
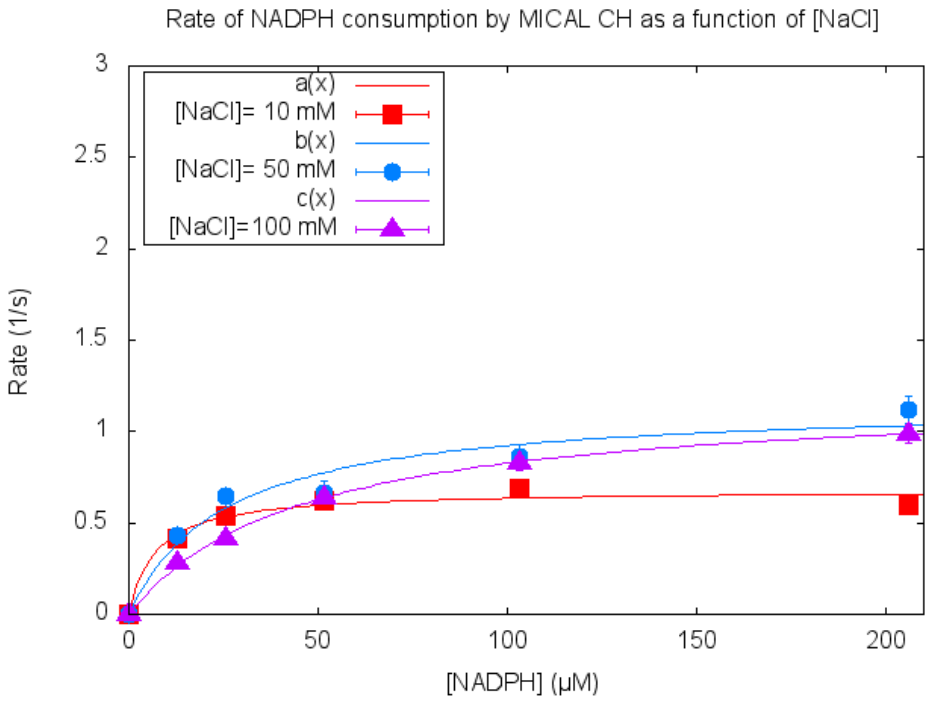


Figure 6. Decrease in H_2O_2 production upon addition of F-actin to MICAL$_{FD-CH}$.

(A) MICAL$_{FD-CH}$ was incubated with NADPH (concentration indicated in figure) and different F-actin concentrations. The decrease in H_2O_2 data were fit with a linear equation and the slope is displayed. (B) Control experiment of MICAL$_{FD-CH}$ and H_2O_2 incubated with different F-actin concentrations. The decrease in H_2O_2 data were fit with a linear equation and the slope is displayed.

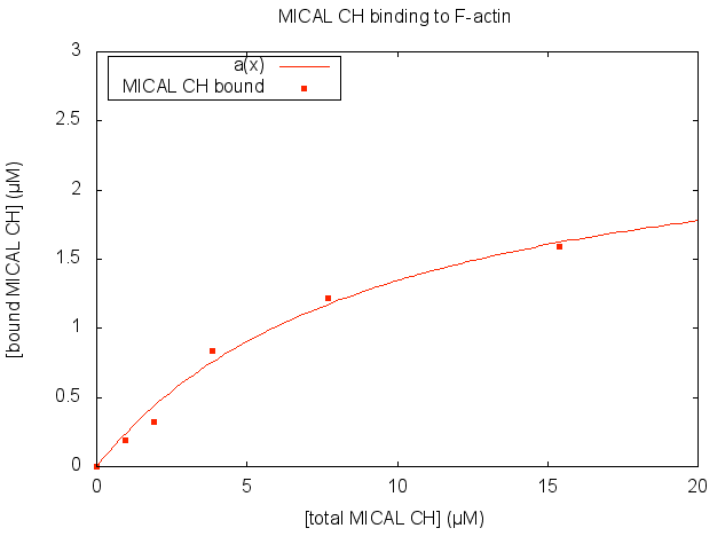


[NaCl] (mM)	K_M (µM)	k_{cat} (s^{-1})
10	7.1 ± 2.4	0.8 ± 0.04
50	25.6 ± 8.1	1.2 ± 0.1
100	46.0 ± 2.6	1.2 ± 0.02

Figure 7. $MICAL_{FD\text{-}CH}$ NADPH consumption at three NaCl concentrations.

Rate of NADPH consumption was measured, for different NADPH concentrations, in the presence of 10 mM (red), 50 mM (purple), and 100 mM (blue) of NaCl. Plot prepared using gnuplot software. The table shows the estimated parameters K_M and k_{cat} for each curve ± standard error.

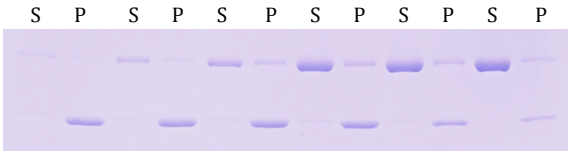
A



K_D = 9.5 ± 2.5 µM

B_{max} = 2.6 ± 0.4 µM

B



[$MICAL_{FD-CH}$] (µM) : 0.95 1.9 3.8 7.7 15.4 30.8

Figure 8. $MICAL_{FD-CH}$ binding to F-actin.

4 µM of actin was first polymerized (see 2.2.3), incubated with 0-31 µM of $MICAL_{FD-CH}$ at room temperature for 1 hr, and ultracentrifuged at 100,000 xg for 30 minutes. Supernatants (S) were carefully removed and added to 2X Laemmlli sample buffer for loading on SDS-PAGE gel. The pellet (P) was resuspended with the actin buffer, incubated at room temperature for 10 minutes, and added to a 2X Laemmlli sample buffer for loading on SDS-PAGE gel. (A) Quantification of the intensities of the $MICAL_{FD-CH}$ bands in the supernatant and the pellet from the SDS-PAGE gel (B) was performed using the ImageJ software [29], and plotted using gnuplot software.

3. Structural Characterization of $MICAL_{FD\text{-}CH}$

The kinetic studies performed with $MICAL_{FD\text{-}CH}$ (chapter 2) suggest that even though the CH domain does not interact with the active site in the FD, it enhances the redox activity in the presence of actin. To gain a more detailed understanding of how this occurs, a major effort in this thesis work was devoted to crystallization and structure determination of $MICAL_{FD\text{-}CH}$. Only with an atomic-resolution structure of the FD and the CH domains contained in a single polypeptide chain can one gain insight into the influence of the CH domain on the activity of the FD. The structure of $MICAL_{FD\text{-}CH}$ reveals that the CH domain contacts the FD far from where NADPH is predicted to bind, in agreement with the interpretation of the kinetics results. Furthermore, it reveals that a flexible linker (disordered in the crystal structure) connects the two domains, and this flexibility may be important for optimizing binding to actin.

3.1 Introduction

Growing axons are guided to their appropriate target tissues by extracellular cues (Figure 1). These guidance cues are essential for appropriate neuronal growth and development, rewiring, fasciculation/defasiculation, and nerve regeneration after injury [3]. Semaphorins, the best characterized class of external repellent guidance molecules, interact with Plexin and neuropilin receptors on axonal growth cones [30, 31]. Upon interaction with extracellular semaphorins, Plexin recruits MICAL (Molecule Interacting with CasL) to its cytosolic domain, and this leads to growth cone collapse as a result of reorganization of the cytoskeleton caused by MICAL activity [5-7].

MICAL is unique in that it combines a catalytic monooxygenase domain with three domains known to be involved in protein-protein interaction, which may be important for recruitment of substrates and modulation of its activity [13]. These include an immediately adjacent CH domain that is usually found in actin-binding proteins. Modulation of MICAL activities by its non-redox domains is crucial since its redox activity can directly oxidize actin filaments leading to their depolymerization, and must be controlled so that it is on or off at the right times.

To gain insight into modulation by the CH domain, steady-state kinetic experiments were initially carried out in the absence and presence of actin (described in chapter 2). The results show that the CH domain could be acting as a substrate specificity factor by enhancing the activity of the FD only when the substrate (actin) is present. The structural determinants for this modulation are

unclear, however. Given that the only available structures are of isolated MICAL domains, the work described in this chapter was dedicated to structural determination and characterization of a fragment containing MICAL's FD and CH domains in a single polypeptide chain by x-ray crystallography.

3.2 Results

3.2.1 Expression and Purification of $MICAL_{FD-CH}$

Cloning

A plasmid containing DNA coding for MICAL-1 from *mouse* (codon-optimized for expression in *Escherichia coli* – Genescript Inc.) residues 2-615 (containing the Q78K mutation to remove an endogenous protease site) was cloned onto a pET28a expression vector containing an N-terminal His-tag with an engineered N-terminal *Tobacco etch* virus (TEV) protease site, and used to transform *Escherichia coli* BL21.

Protein expression and purification

After induction by addition of 0.2 mM β-D-1-thiogalactopyranoside (IPTG), cells were grown for 15 hours at 17°C in LB media before harvesting by centrifugation and freezing the pellets in -80˚C. Frozen pellets were resuspend in lysis buffer containing 50 mM Tris-HCl pH 7.0, 140 mM NaCl, 10 mM Imidazole, 0.1 % Tween-20, 5 mM $MgCl_2$, 2 mM β-mercaptoethanol (BME), 5 mM Benzamidine, and 10 % v/v glycerol. Cells were broken via microfluidization, and the lysate was centrifuged at 23,000 xg for 60 minutes. The supernatant was collected and filtered in a 0.22 micron Stericup filter (Millipore) before loading on an AKTA FPLC machine. The protein was purified by Ni-Sepharose affinity chromatography (His trap column, 5 ml, GE) with a gradient of 50 mM Tris-HCl pH 7.0, 140 mM NaCl, 10-

500 mM imidazole. The eluted protein was dialyzed against 50 mM Tris-HCl pH 7.0, and 200 mM Glycine, and digested overnight with TEV protease to remove the His-tag. The cleaved product was purified by using a Source 15S cation exchange column, eluting with 50 mM Tris-HCl pH 7.0, and 0-1 M NaCl gradient (Figure 9). The final yield was ~ 3 mg per liter of culture. After dialysis to buffer exchange to 50 mM Tris-HCl pH 7.0, 200 mM NaCl, $MICAL_{FD\text{-}CH}$ was concentrated to ~ 25 mg/ml using a 3 kDa MWCO ultrafiltration device (GE Healthcare), before storage at -80˚C. $MICAL_{FD\text{-}CH}$ is unstable without the presence of salt and attempting to concentrate the protein in Tris buffer lacking NaCl causes precipitation of the protein.

3.2.2 Crystallization of $MICAL_{FD\text{-}CH}$

Differential scanning fluorimetry

Before initiating crystallization trials, the thermal stability of purified $MICAL_{FD\text{-}CH}$ was subject to differential scanning fluorimetry analysis [32]. This experiment measures the thermal stability of $MICAL_{FD\text{-}CH}$ in varying buffer, pH, and salt conditions. Increased thermal stability of the protein could enhance the probability of crystallization. Results are shown in a qualitative manner using the ThermoQ software (Figure 10).

Crystallization trials

From the results of the differential scanning fluorimetry, I decided to use MES pH 6.0 buffer for protein storage prior to screening for crystallization

conditions. Initial screening attempts (1 µl of protein solution combined with 1 µl of reservoir solution and equilibrated against 500 µl of reservoir solution) were based on screening around the crystallization condition for $MICAL_{FD}$ (PDB 2BRA), containing PEG 2000 MME and sodium acetate pH 4.6, in a hanging-drop 24-well format. Some phase separation and microcrystals were observed, but nothing suitable for data collection. Trials in these same conditions with the protein being in sodium citrate buffer pH 5.0 instead of MES buffer pH 6.0 gave the first "proto-crystals" of $MICAL_{FD\text{-}CH}$ (Figure 11). Sodium citrate buffer pH 5.0 also enhanced thermal stability of $MICAL_{FD\text{-}CH}$ similar to the MES buffer. The first sign of "needles" was observed by screening with the same conditions [21 % PEG 2000 MME, 0.1 M sodium acetate pH 5.0] but in a sitting-drop vapor diffusion setup as opposed to hanging drop (Figure 11).

Final conditions and microseeding

The "needles" were optimized to a "large cluster of needles" by using HEPES buffer pH 7.5 was used instead of sodium acetate pH 5.0 in the crystallization condition [20-22 % PEG 2000 MME, 0.1 M HEPES pH 7.5] (Figure 11). This cluster of needles was used for microseeding, varying the dilution of the seeds in the mother liquor, until "cluster of rods" grew where single rod-shaped crystals could be picked and mounted for data collection. The "large cluster of needles" is not reproducible, and to grow more $MICAL_{FD\text{-}CH}$ crystals microseeding from a drop that contains some needles/rods is required.

In summary, for crystallization of $MICAL_{FD\text{-}CH}$, 1 µl of protein solution [25 mg/ml in 100 mM sodium citrate pH 5.0, 200 mM NaCl] was combined with 1 µl of reservoir solution [100 mM HEPES pH 7.0, 20 % PEG 2000 MME] containing microseeds and equilibrated against 500 µl of reservoir solution. Crystals grew in 24 hours at 20°C.

3.2.3 Structure Determination of $MICAL_{FD\text{-}CH}$

3.2.3.1 Native data collection from one crystal (crystal 1)

For native data collection, a single rod in a bunch of rods was isolated and mounted on a nylon loop (Hampton Research). Prior to mounting, the thin film that forms on the surface of the drop containing the crystals needed to be removed gently with a loop or needle. The crystal was washed in a drop containing the crystallization condition, and once more in a similar drop but with 10 % glycerol added for cryoprotection. Crystals were flash-frozen in liquid nitrogen.

$MICAL_{FD\text{-}CH}$ crystals diffracted up to moderate 2.9 Å resolution on a Saturn 944+ CCD detector (Rigaku Americas Corporation, The Woodlands, TX) with in-house light source FR-E+ Super-Bright™ generator (Rigaku Americas Corporation, The Woodlands, TX) equipped with VariMAX mirrors and a copper rotating anode generating X-rays with wavelength of 1.5418 Å.

Crystals belong to the $P2_1$ space group with cell dimensions: a = 64.5 Å, b = 49.9 Å, c = 94.7 Å, , α = = 90.0 °, β = 107.9 °. They contained one $MICAL_{FD\text{-}CH}$

molecule per asymmetric unit. Native data were reduced and processed with HKL2000 (HKL Research Inc.). Data collection statistics are shown in Table 3.1.

3.2.3.2 Phase Determination

To obtain the three-dimensional structure of the protein from the diffraction data, phase angles must be determined (see chapter 5 for background of the technique). $MICAL_{FD\text{-}CH}$ spans the first 615 amino acids of the MICAL protein, and since the crystal structure of $MICAL_{FD}$ (PDB 2BRA) previously determined in our laboratory spans the first 484 amino acids i.e. ~ 80 % and is the exact same FD, the phases were determined by the molecular replacement method with the program MOLREP (CCP4 suite) [33, 34], using the crystal structure of $MICAL_{FD}$ as a search model (PDB 2BRA) [16].

3.2.3.3 Model Building and Refinement

The program COOT [35] was used for model building, and subsequent rounds of refinement were performed in the program REFMAC (CCP4 suite) [33, 34]. The residues of the FD required little refinement since $MICAL_{FD}$ was used as the search model. A loop (amino acids 147-150) was missing from the $MICAL_{FD}$ crystal structure was since there was no density observed for it; this loop was modeled in the current $MICAL_{FD\text{-}CH}$ crystal structure (Figure 12). The residues of the CH domain were manually built using the difference Fourier map. Initially, poly-alanine alpha helices were traced into well-defined regions of the electron density. Then, the NMR

model of *human* MICAL's CH domain (PDB 1WYL) was placed using the initially built helices as a guide, followed by several rounds of refinement in REFMAC [33]. Refinement statistics are shown in Table 3.1.

The structure from crystal 1 shows electron density for both the FD (residues 7-484) and CH (residues 511-611) domains. However, there was no observable electron density for residues 485-510 (the linker between the two domains). This introduced an uncertainty about which CH domain among three symmetry mates belonged to the same molecule with given FD. Upon analysis of the asymmetric unit, it appeared that the 25-residue linker could connect the FD to either of three possible choices of a CH domain in the crystal asymmetric unit (Figure 13). The first step to resolve this ambiguity was to ensure that $MICAL_{FD\text{-}CH}$ had crystallized as a single polypeptide chain, and that no cleavage had occurred at any stage after purification and before crystallization. SDS-PAGE analysis of re-dissolved crystals confirms that $MICAL_{FD\text{-}CH}$ crystallized as a single polypeptide chain (Figure 14).

3.2.3.4 Native data collection from a 2nd crystal (Crystal 2)

X-ray diffraction data from a second crystal was essential for resolving the ambiguity in the FD and CH domain connectivity. Data was collected as described for crystal 1, with the only difference being that the $MICAL_{FD\text{-}CH}$ model from crystal 1 was used as the molecular replacement search model for phase determination. The model was built and refined in the same way as for crystal 1. Even in this crystal, no electron density was observed for the linker between the two domains.

3.2.3.5 Comparison of crystal 1 and crystal 2 unit cells

$MICAL_{FD\text{-}CH}$ crystallized in two different crystal forms sharing the same space group with slightly different cell dimensions. In both crystal forms, the unit cell is made by three layers stacking in the direction of the cell long axis, with two layers formed exclusively by FD domains that sandwiches a layer of CH domains (Figure 15). The layer of FD domains makes the major crystal contacts across unit cells and minor ones across the CH layer. Each FD in the asymmetric unit interacts with three (four in crystal form 2) CH domains (option 1, 2, 3).

Examination of the asymmetric unit in crystal 1 and 2 shows that one FD can connect to one of three possible CH domains. Superposition of the FD and its respective CH domain connectivity options from crystal 1 onto that of crystal 2, followed by analysis of RMSD between the CH domain options from both crystals clarified the connectivity (Figure 13). The RMSD (Root Mean Square Deviation) of option 1 and 2 from both crystals is 3.7 Å and 3.8 Å, respectively, while the RMSD for option 3 is significantly lower at 1.6 Å. This suggests that option 3 shows the true connectivity, as it is the most similar between the two crystals.

Table 1. MICAL$_{FD-CH}$ data collection and refinement statistics.

Data set	Native #1	Native #2
Data collection		
Space Group	$P2_1$	$P2_1$
Cell dimensions (a, b, c)	64.5, 49.9, 94.7	70.2, 50.2, 97.0
Cell angles (α, β,), °	90, 107.9, 90	90, 101.2, 90
Resolution, Å (Last shell)	90-2.9	95-2.9
Total reflections (unique)	45672 (11353)	52018 (14602)
Completeness, % (Last shell)	83.1 (20.7)	96.2 (65.1)
I/σ (Last shell)	9.3 (1.4)	6.6 (1.1)
Rsym (Last shell)	0.12 (0.36)	0.18 (0.64)
Refinement		
$R_{crystal}/R_{free}$	0.27/0.30	0.23/0.30
Stereochemistry		
rms bond length, Å	0.004	0.006
rms angles, °	0.8	1.2
Temp factors, Å^2		
B-factor protein	19.4	29.6
B-factor FAD	13.7	15.9

3.2.3.6 SAXS analysis of $MICAL_{FD-CH}$

To obtain further insight into the behavior of $MICAL_{FD-CH}$ in solution and compare it with the model from the crystals, we performed SAXS (Small-angle x-ray scattering) analysis on $MICAL_{FD-CH}$. The overall SAXS parameters are presented in Table 3.2. At concentrations greater than 3 mg/ml $MICAL_{FD-CH}$ is in equilibrium between a monomer and a dimer (option 3), as evident by fitting the theoretical scattering curves of an ensemble of monomer (~90 %) and dimer (10 %) to the experimental data compared with monomer alone (Figure 16) [36]. Furthermore, CH domain connectivity option 1 and option 2 fit the experimental data better than option 3. Option 3 has a lower estimated Rg of ~24 Å than option 1 and 2 with an estimated Rg of ~29 Å, with the latter options matching the Rg calculated from the experimental SAXS scattering curves. This suggests that the disordered linker allows for some flexibility of the CH relative to the FD.

Table 2. SAXS parameters of $MICAL_{FD-CH}$.

Concentration (mg/ml)	I(0)	Rg (Å)	V_p (Å^3)
1	27.5	28.7	102,000
2	56.4	29.3	114,000
3	90.6	30.0	122,000
7	414.0	33.3	116,000

3.2.4 Structure of $MICAL_{FD\text{-}CH}$

The structure of $MICAL_{FD\text{-}CH}$ contains the FD and CH domains of *mouse* MICAL-1 in the same polypeptide chain. $MICAL_{FD\text{-}CH}$ crystals contain one molecule in the asymmetric unit (Figure 17).

The linker connecting the FD and CH domains

The FD and CH domains are connected by a 25-residue linker, which has no observable electron density. This disordered linker (amino acids 485-510) is not very well conserved amongst different MICALs. Based on its sequence, it is predicted to be intrinsically unstructured by secondary structure prediction algorithms, for example PSIPRED server (Expasy).

Chromatographic behavior

Size exclusion chromatography, using a solution containing 7 mg/ml (~100 μM) of $MICAL_{FD\text{-}CH}$, shows that $MICAL_{FD\text{-}CH}$ is monodisperse and behaves as a monomer (Figure 18), with an estimated molecular weight of ~63 kDa, in good agreement with the expected molecular weight of 68 kDa.

Sequence conservation

From the N-terminal FD (amino acids 1-484) to the end of the CH domain (amino acids 511-611), MICAL-1 is highly conserved (Figure 19) from *Drosophila* to *humans*. With regards to *humans*, $MICAL_{FD\text{-}CH}$ from *mouse* used in this work is 91%

identical and 93% similar to the $MICAL_{FD\text{-}CH}$ region of MICAL-1 in *humans*. Therefore, it is very likely that the structure of a $MICAL_{FD\text{-}CH}$ construct from *humans* is similar to that from *mouse*.

Analysis of the interface contact between the FD and CH domains

Residues involved in contacts between the FD and CH domains include Met 232, Lys 235, Ile 264, Met 369, Leu 402, Arg 408, Gln 441, Leu 442, Ser 444, Gln 445, Ser 447, Glu 449, and Asn 450 from the FD. The aforementioned residues make contacts with residues Leu 553, Leu 565, Thr 569, Arg 573, Val 574, Glu 576, His 577, Glu 578, Gly 580, Thr 582, Pro 583, Val 584, Ser 586, Gln 588, and Ala 609 of the CH domain (Figure 20). A salt bridge is formed between Arg 408 and Glu 576 side chains, and both these residues are conserved from *Drosophila* to *human*. A hydrophobic contact between Ile 264 and Ala 609 is also present. The amide nitrogen of Lys 235 makes a polar contact with the side chain of Glu 578, as does the side chain of Ser 447 and the carbonyl carbon of Pro 583.

The interface has a solvent buried surface area of ~900 $Å^3$, and having this large of a buried surface area is highly unlikely to be due to crystal contacts, further supporting our current interpretation of option 3 being the CH belonging to the asymmetric unit.

To test whether it is feasible to connect the FD and CH domains via the 25-amino acid linker, the linker was modeled using MODELLER [37]. It is possible for the linker to connect the two domains, however, it is in a very extended conformation. This raises the possibility that the connectivity observed in the

crystal is a result of domain swapping; where the CH domain (option 3) belongs to a neighboring FD. Based on SAXS analysis, $MICAL_{FD-CH}$ is in equilibrium between monomer and dimer at high concentrations (e.g. in a crystal). One important point to note is that signaling proteins are typically found between 10 nM and 1 μM inside the cytosol of a cell, nowhere near the concentration observed for oligomerization of $MICAL_{FD-CH}$.

3.2.4.1 Structure of the FD

The FD is a monooxygenase domain that consists of a mix of α-helices and β-sheets. Both parallel and anti-parallel β-sheets can be found. Another structural feature present in the FD is a Rossman β-α-β fold that contains a GXGXXG motif, a sequence common in oxidoreductase enzymes dependent on NADPH and FAD cofactors. The FAD cofactor in the active site that is necessary for the enzymatic activity of MICAL [16] is oxidized, and is solvent-accessible through a large opening in the active site. It is highly similar to the structures of the isolated FD (PDB 2BRA, 2BRY), with an rms deviation of 0.5 Å (Figure 21) [16, 20].

3.2.4.2 Structure of the CH domain

Structural description of the CH domain

The CH domain of $MICAL_{FD-CH}$ shares the same architecture as other CH domains of the same type found in actin-binding proteins: three parallel tightly packed α-helices, with a fourth α-helix perpendicular to the other three. Key

hydrophobic residues in these helices that are important for helix packing and binding F-actin are also conserved (Figure 22). Structurally, the CH domain of $MICAL_{FD\text{-}CH}$ construct is a type-2 CH, highly similar to the human MICAL-1 CH domain (PDB 1WYL), observed in the structure determined by NMR [21]; the structures align with a deviation of 0.8 Å rms (Figure 23).

Similarity to CH domains in actin-binding proteins

The CH domain of $MICAL_{FD\text{-}CH}$ shows conserved structural features found in the type-2 CH domains of actin-binding proteins, such as conserved residues of actin-binding found in the helix perpendicular to the other three, and a conserved PIP_2 binding site (Figure 22, 23) [18].

3.2.5 Structure-function implications

What is immediately evident upon inspection of the structure is that the CH domain in $MICAL_{FD\text{-}CH}$ is not located close to the large opening of the active site of the FD. Thus, the CH domain does not interfere with accessibility to the active site of the FD. This is consistent with kinetics data (Chapter 2) that show that $MICAL_{FD}$ and $MICAL_{FD\text{-}CH}$ catalyze the oxidation of NADPH by O_2 with similar kinetics.

Furthermore, the presence of the CH domain does not significantly alter the conformation of the active site in the FD. The FD in $MICAL_{FD\text{-}CH}$ aligns with a 0.5 Å rms deviation (422 residues aligned) to the $MICAL_{FD}$ crystal structures (PDB 2BRA, 2BRY). This is consistent with the modular nature of CH domains and the fact that

the CH domain in $MICAL_{FD\text{-}CH}$ shares the same fold and architecture as the isolated CH domain from *human* MICAL (PDB 1WYL).

The CH domain of MICAL is close to where O_2 is predicted to enter the FD active site near a small channel that opens upon reduction of the FAD by NADPH [13, 16, 20]. The conserved residues that form the actin-binding segment, based on sequence alignment and studies on CH domains in actin-binding proteins, are surface exposed and, relative to the FD, oriented in such a way that could bind actin filaments and optimize the interaction between actin's Met 44 (residue that is oxidized in actin) and the FAD cofactor of MICAL. This optimization of the binding to F-actin may explain the kinetic data that shows an observed *increase* in redox activity of $MICAL_{FD\text{-}CH}$ *in the presence of F-actin* compared with $MICAL_{FD}$. In addition to $MICAL_{FD\text{-}CH}$ being able to oligomerize at increasing concentrations, the fact that the linker between the FD and CH domains is disordered and does not adopt a single conformation may help mediate this accommodation of the binding to F-actin.

3.3 Conclusions and Discussion

The importance of the CH domain of MICAL has been highlighted by the fact that it is highly conserved across all organisms possessing MICAL proteins [13], and by experiments in *Drosophila* showing that axon guidance and MICAL-mediated actin modification requires the presence of the CH domain *in vivo* [5-7]. Furthermore, even though it was previously known that CH domains are usually involved in actin-binding, the way in which this occurs for MICAL was not known. The structure of $MICAL_{FD\text{-}CH}$ presented in this thesis work provides valuable insight into the function of the CH domain.

The structure is in agreement with kinetics data (Chapter 2) that show that $MICAL_{FD}$ and $MICAL_{FD\text{-}CH}$ catalyze the oxidation of NADPH by O_2 with similar kinetics. This also supports a model of NADPH binding to the large opening of the active site in the FD, based on observed electron density for Cl^- ion in the $MICAL_{FD}$ crystal structure (PDB 2BRA) [16]. In this model, the phosphate of NADPH is modeled occupying the position where the Cl^- ion was observed (Figure 24).

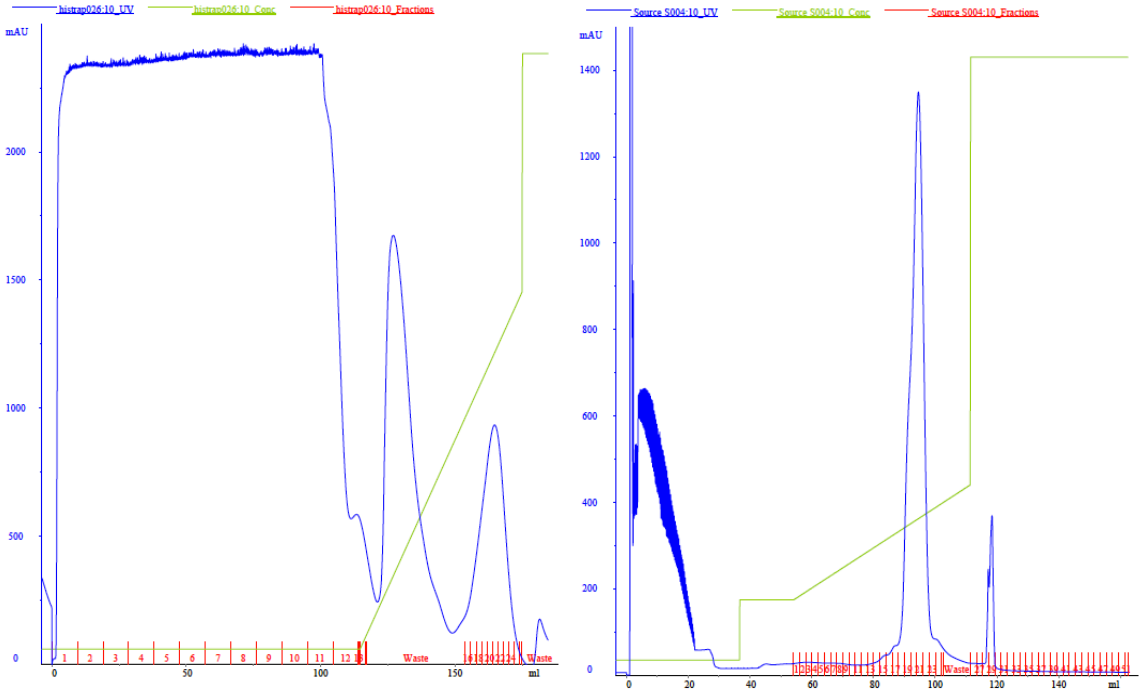
In the actin-binding domains of proteins that cross-link actin filaments (e.g. α-actinin, spectrin, dystrophin), the two tandem CH domains (type-1 and type-2) bind F-actin with high affinity, and while CH type-1 can bind F-actin alone, CH type-2 alone cannot bind F-actin [18]. Furthermore, the structural data on the CH domains of these actin-binding proteins suggests that even though they are very similar structurally, they do not bind F-actin in a similar manner [38, 39]. The CH domain of MICAL (type-2) may be functioning in a similar way, where the FD

functions as a type-1 CH domain in that it can bind F-actin alone [6], and the presence of the CH domain may increase the binding of the FD to F-actin. Actinin, utrophin, and dystrophin bind F-actin in an open conformation whereas the free proteins exist in a closed conformation. The ability of the two tandem CH domains to undergo some conformational change with respect to each other upon binding F-actin is important, as mutations in the interface between the two tandem domains that affects this result in disease phenotypes brought about by disregulation of the F-actin binding activity [39]. Similarly, the disorder of the loop between the FD and CH domains of MICAL may allow the two domains to accommodate for optimal binding to F-actin.

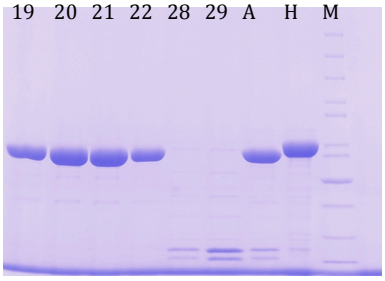
With kinetics and structural inferences taken together, we propose a model of how the CH modulates the activity of MICAL in the presence of F-actin (Figure 25, detailed explanation in Chapter 4). In this model, NADPH can enter the active site from the large opening (as modeled in Figure 24), and Met 44 of actin can be accommodated through the small channel that opens up on the other side. Kinetic data of $MICAL_{FD\text{-}CH}$ redox activity as a function of increasing F-actin suggests that the K_M of NADPH binding does not change; only the k_{cat} is increased, further supporting this model. Met 44 of actin, the residue modified by MICAL, is contained in the D-loop of actin (residues 39-52). It was shown experimentally that the D-loop of actin is flexible and that in solution it can adopt multiple conformations [40]. It is disordered in most crystal structures of actin, and does not adopt the same conformation in the structures in which it was observed. It is located at the pointed end of the actin monomer, and makes extensive electrostatic and hydrophobic

contacts with the adjacent monomer in the F-actin filament [41]. Thus, the conformation and integrity of the D-loop is important for filament stability [42]. The flexibility of the D-loop may also be important for allowing it to bind optimally to MICAL for oxidation by the reduced FAD.

A



B



Fractions 19-29 from cation exchange chromatogram (above, right)

A: after TEV cleavage, before cation exchange

H: fractions 16-24 from nickel affinity chromatography purification (above, left)

Figure 9. Purification of MICAL$_{FD-CH}$.

(A) Purification of MICAL$_{FD-CH}$ with nickel affinity chromatography and cation exchange chromatography. Left: chromatogram from a Histrap™ HP column. Right: chromatogram from a SOURCE15S column. (B) SDS-PAGE analysis of the purification.

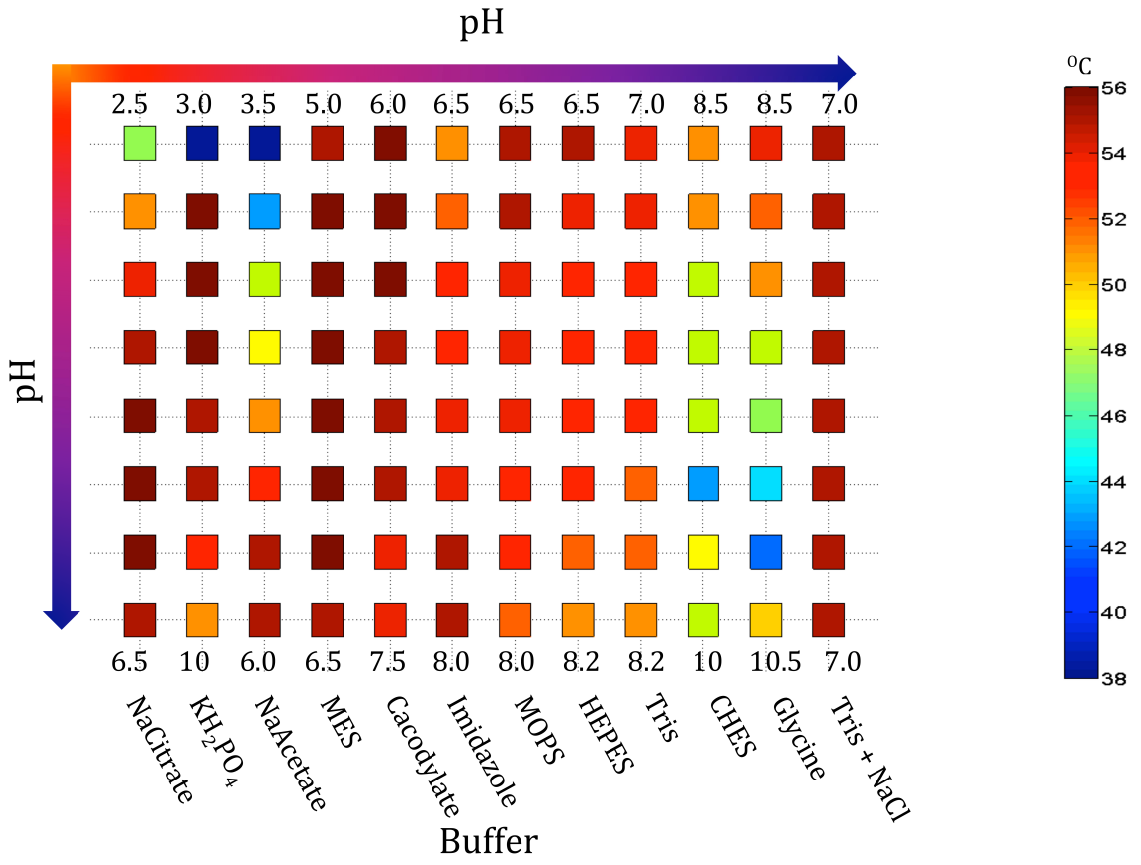
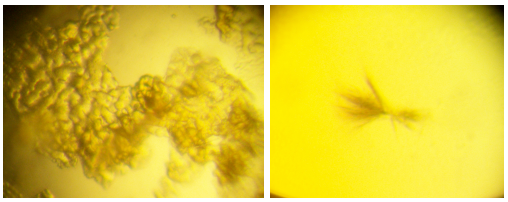


Figure 10. Differential scanning fluorimetry.

Figure depicts the conditions in which the thermal stability of $MICAL_{FD\text{-}CH}$ was tested. Colors indicate the melting point T_m, based on the scale on the right hand side. Higher T_m (more red) indicates higher thermal stability.

A



B

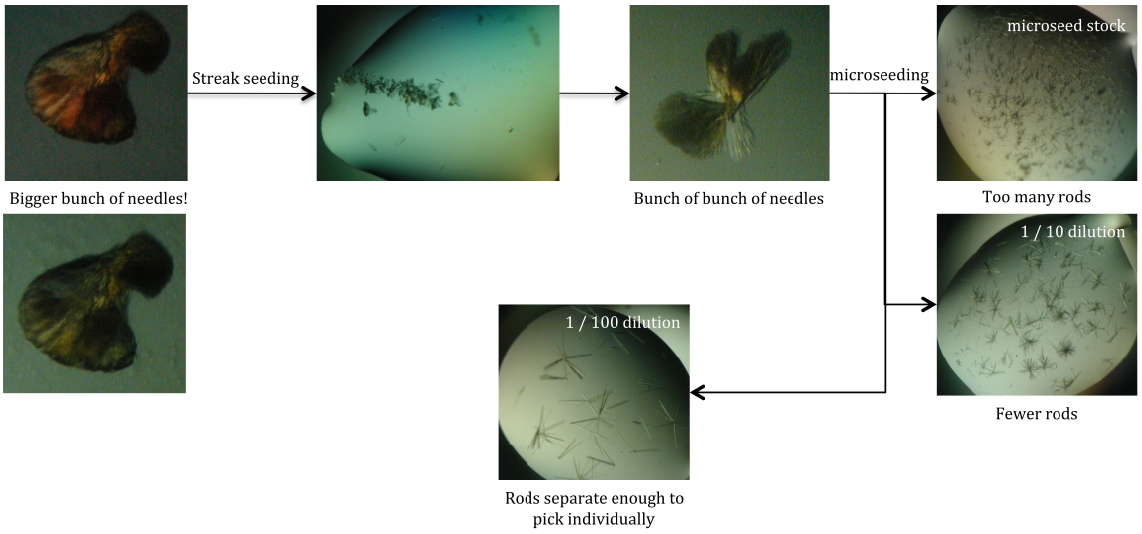


Figure 11. Crystallization of MICAL$_{FD-CH}$.

(A) Left: first "proto-crystals" observed. Right: first drop containing needles observed. (B) Flowchart of optimization of the crystals, from "large cluster of needles" (photos on the left) to the rods (bottom).

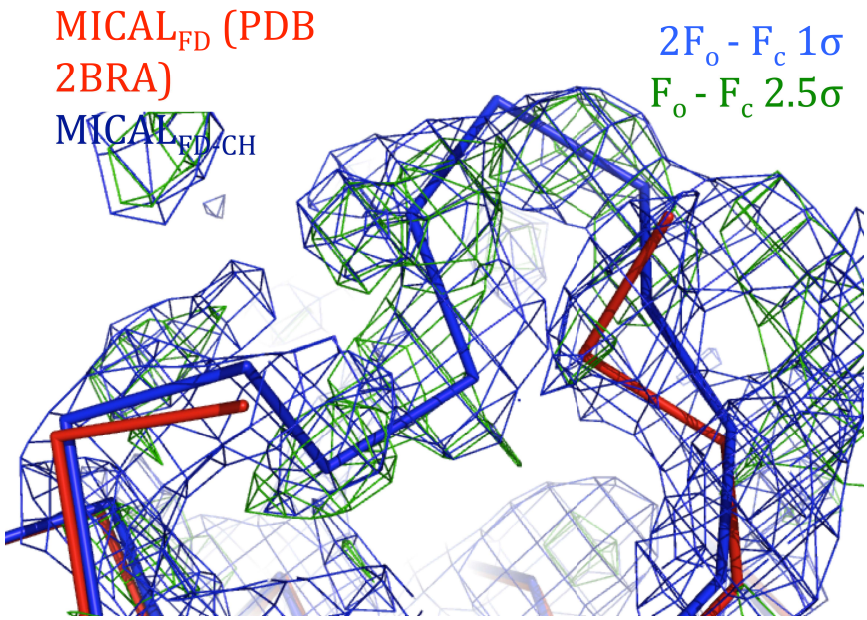


Figure 12. Modeling missing loop of the FD in MICAL$_{FD-CH}$.

During model building of the crystal structure of MICAL$_{FD}$ (PDB 2BRA), no electron density was observed for residues 146-150. They were modeled in MICAL$_{FD-CH}$ using the omit map.

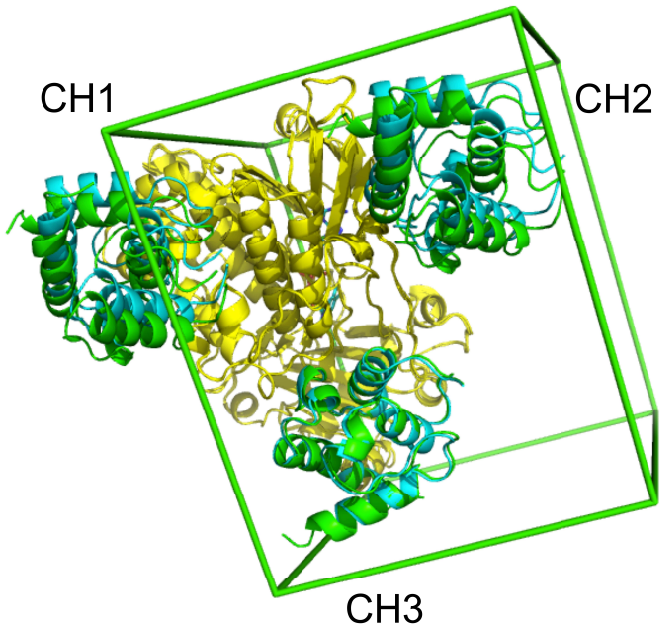


Figure 13. Comparison of connectivity options of MICAL$_{FD\text{-}CH}$.

Options from crystal 1 are in cyan, and options from crystal 2 are in green. Square represents the asymmetric unit. Option 3 (CH3) shows the lowest rms of 1.56 Å, whereas options 1 and 2 have an rms of 3.7 Å between both crystal forms.

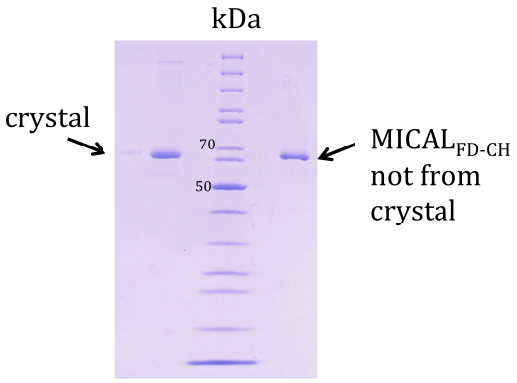


Figure 14. SDS-PAGE analysis of MICAL$_{FD-CH}$ crystal.

Molecular weight marker is the "kDa" lane. Sample from crystal and a sample from a purified solution of MICAL$_{FD-CH}$ are indicated.

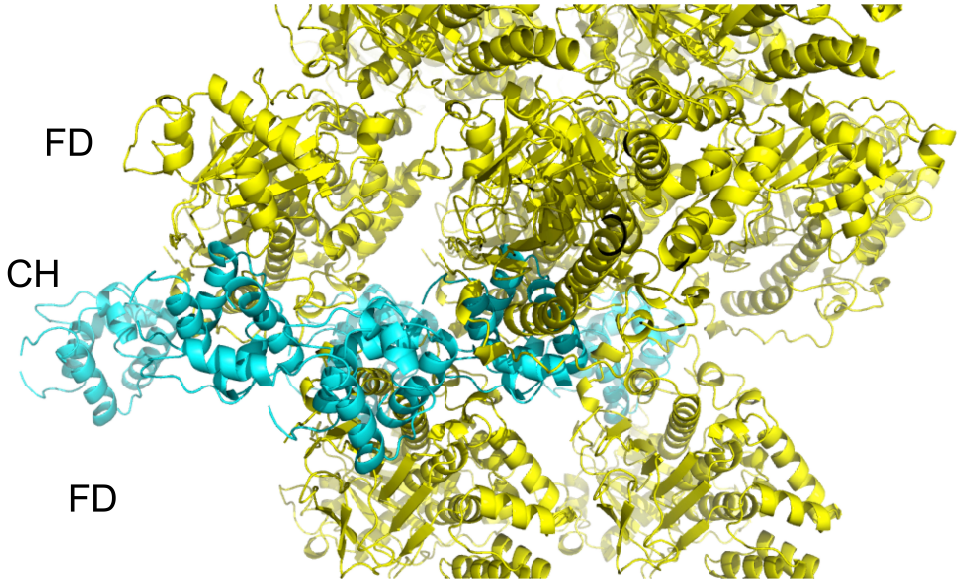
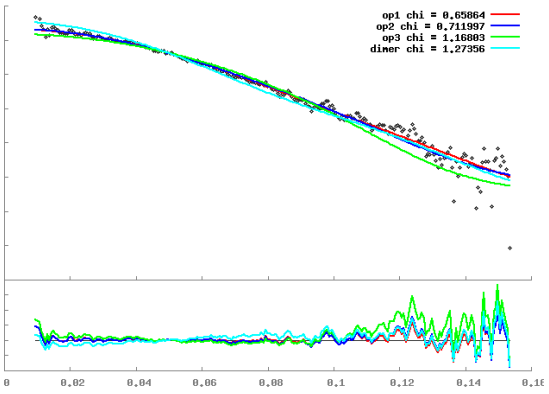


Figure 15. Examination of the crystal packing from MICAL$_{FD-CH}$ crystals.

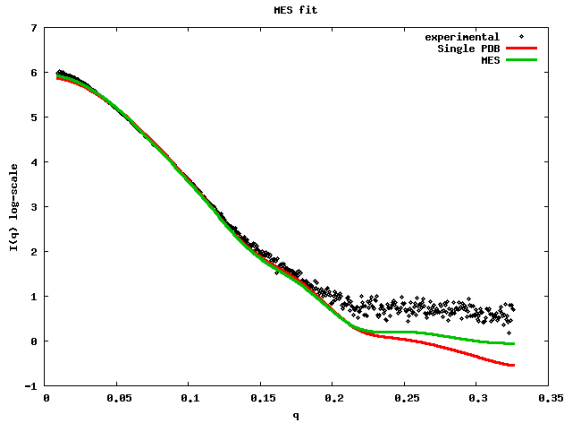
The FD (yellow) and CH (cyan) domains are indicated.

A



PDB file	χ	R_g
op1	0.66	28.17
op2	0.71	29.01
op3	1.17	24.52
dimer	1.27	35.62

B



MES Fit	**χ = 2.033499**
File name	Weight
dimer	0.17148880
op2	0.82820862
Single PDB Best Fit	**χ = 2.38606**
op2	1.00

Figure 16. Comparison of theoretical and experimental scattering from SAXS.

Experimental data is indicated in black filled circles, and theoretical scattering fit through the experimental data is indicated by solid colored line. Op1: option 1; op2: option 2; op3; option 3; dimer: domain-swapped dimer of option 3. (A) Fit for low concentration (merged dataset of $MICAL_{FD\text{-}CH}$ 1, 2, 3 mg/ml). Table shows the theoretical Rg for each option. (B) Multiple ensemble search (MES) fit compared with a single pdb fit to data of $MICAL_{FD\text{-}CH}$ 7 mg/ml. Table shows weights used in the ensemble search.

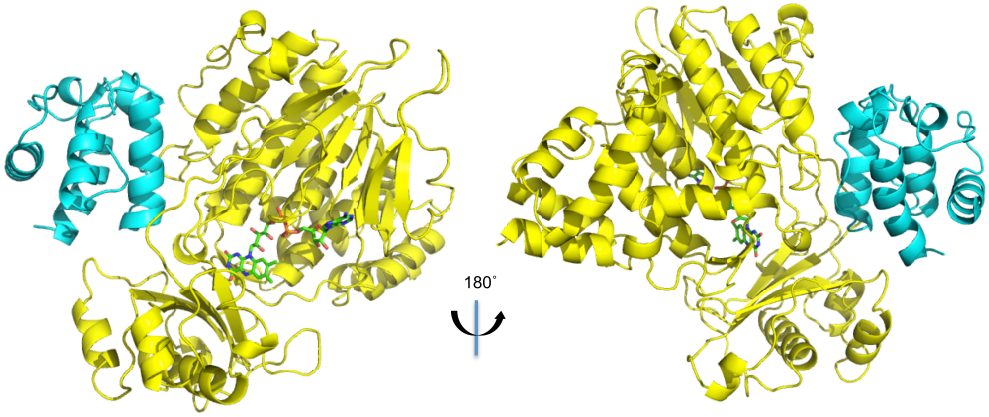


Figure 17. Structural overview of MICAL$_{FD-CH}$.

FD is shown in yellow, CH domain in cyan, and FAD cofactor shown in green sticks.

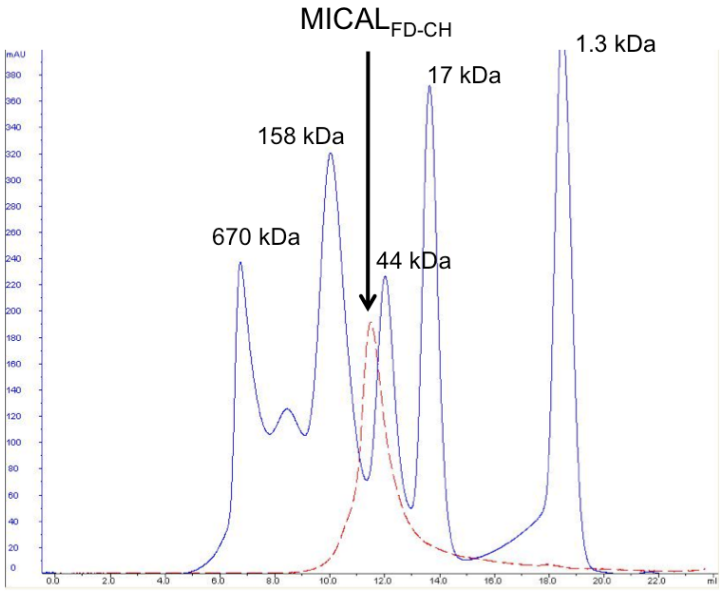


Figure 18. Size exclusion chromatography of MICAL$_{FD-CH}$.

MICAL$_{FD-CH}$ eluted as a monomer (dashed lines, red) in size exclusion chromatography experiment using a Superose-12 (GE) column. Standards are shown in blue, with sizes indicated.

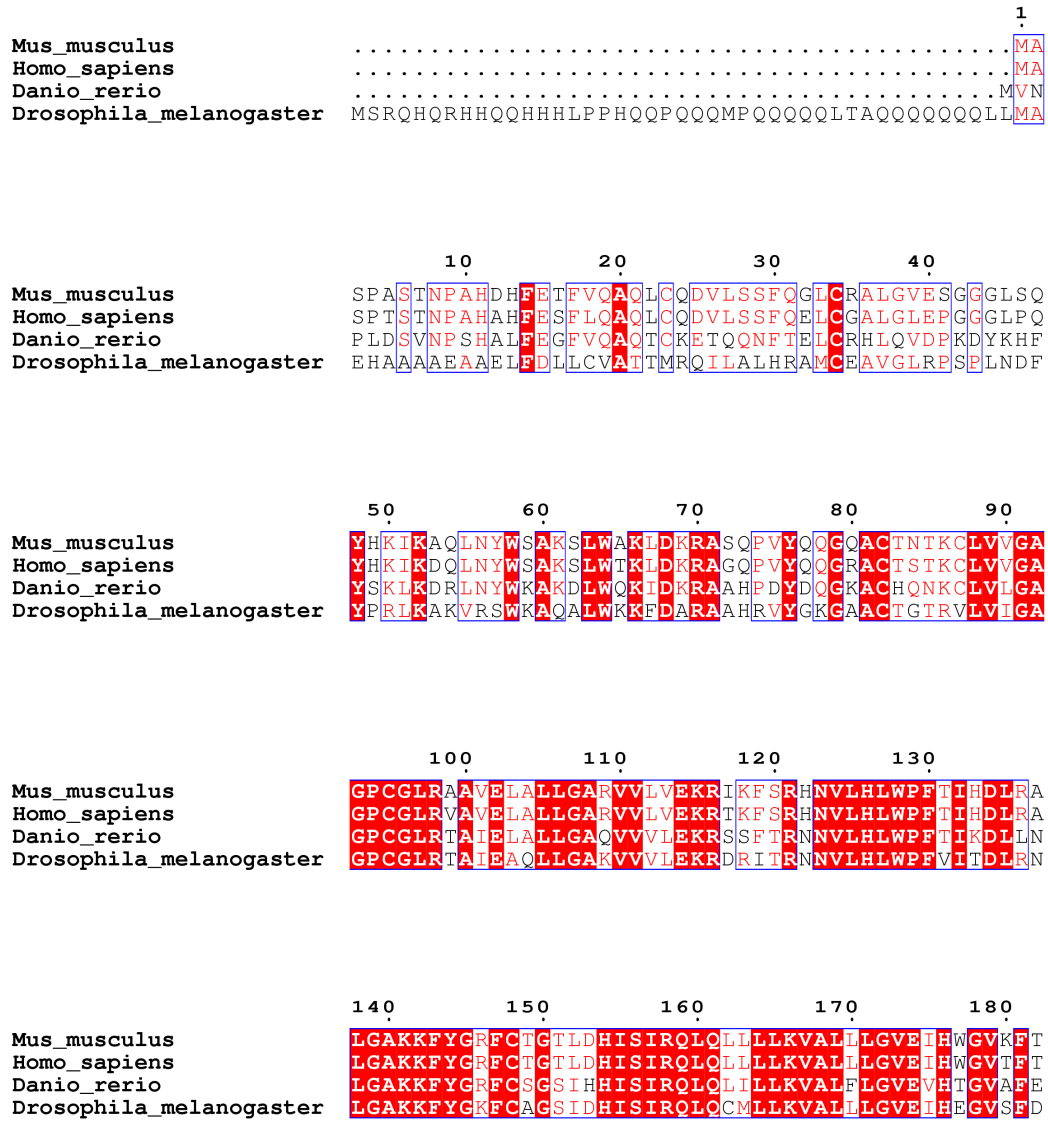


Figure 19. Sequence alignment of MICAL-1 homologues.

MICAL sequences from the indicated organisms were aligned from the N-terminus to the end of the CH domain (aa 1-615). Alignment was performed using ClustalW. Figure prepared with ESPript. Interface residues underlined with an arrowheads.

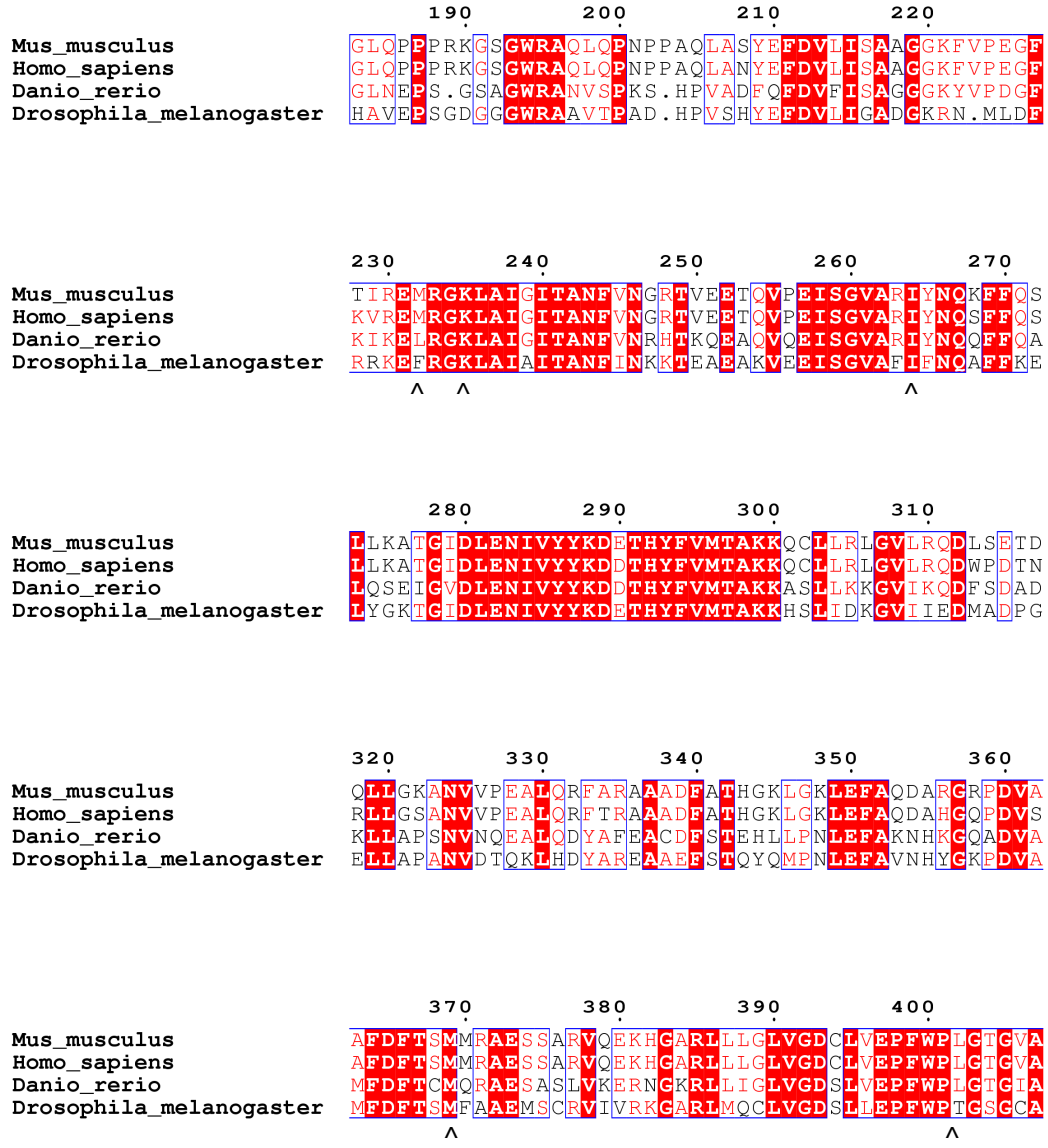


Figure 19. Sequence alignment of MICAL-1 homologues (Cont.).

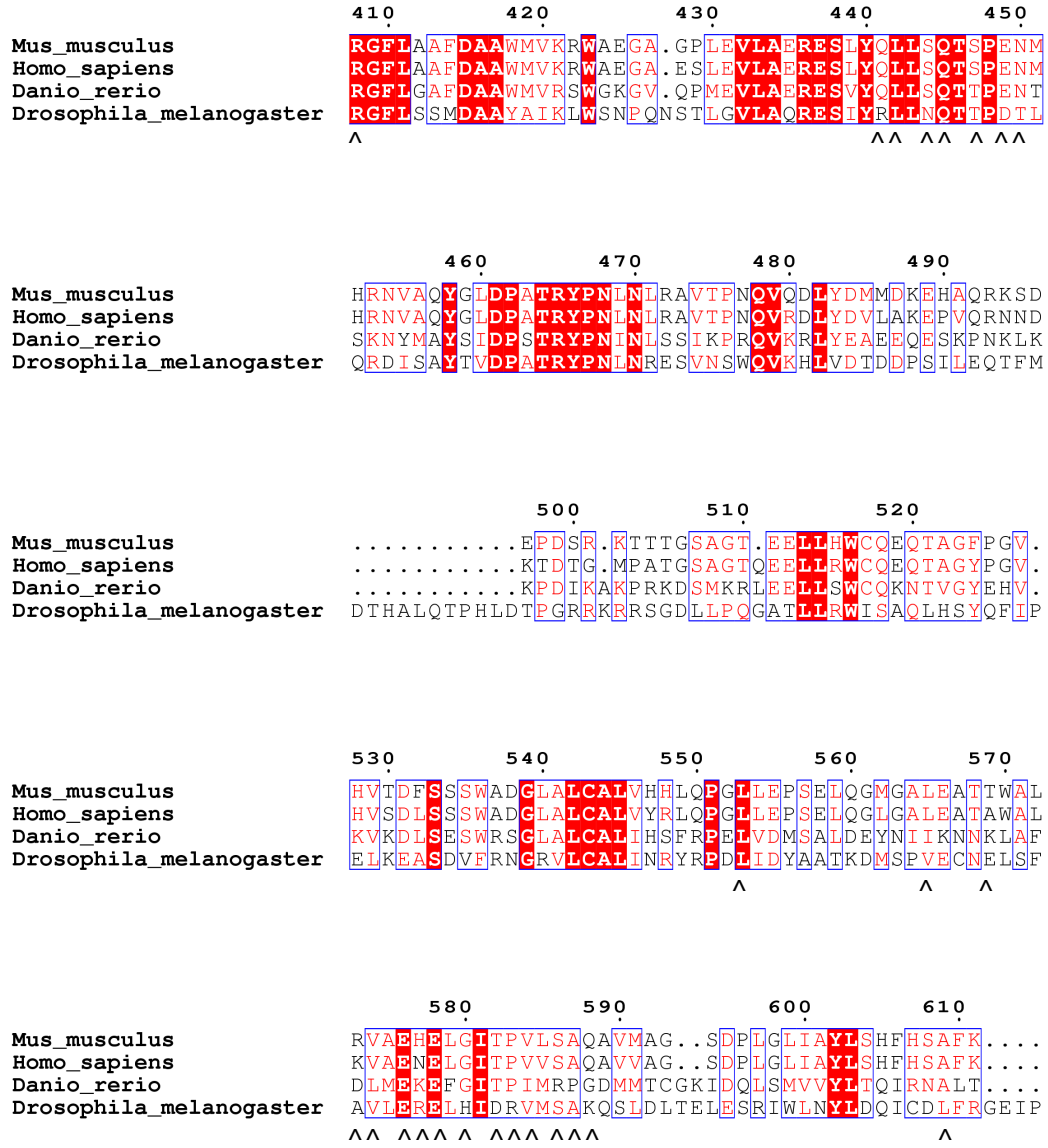


Figure 19. Sequence alignment of MICAL-1 homologues (Cont.).

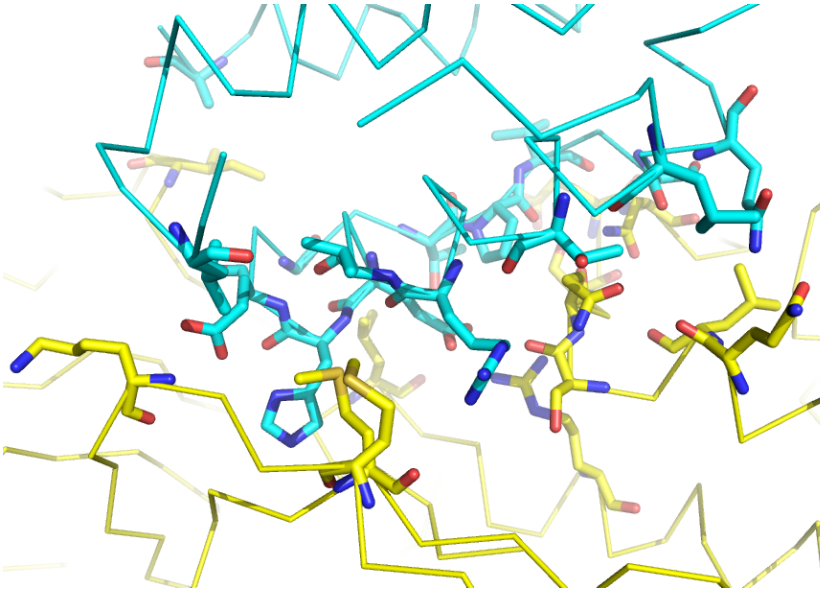


Figure 20. Overview of the interface between the FD and CH domains.

FD is shown in yellow, and the CH domain is shown in cyan. Interface residues are shown in sticks.

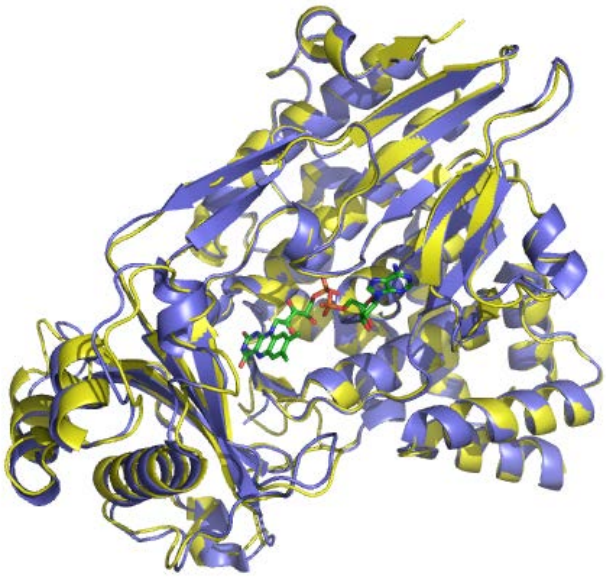


Figure 21. Structural alignment of the FD.

The FD of $MICAL_{FD\text{-}CH}$ (yellow) and $MICAL_{FD}$ (PDB 2BRA, blue) align with an rms of 0.5 Å.

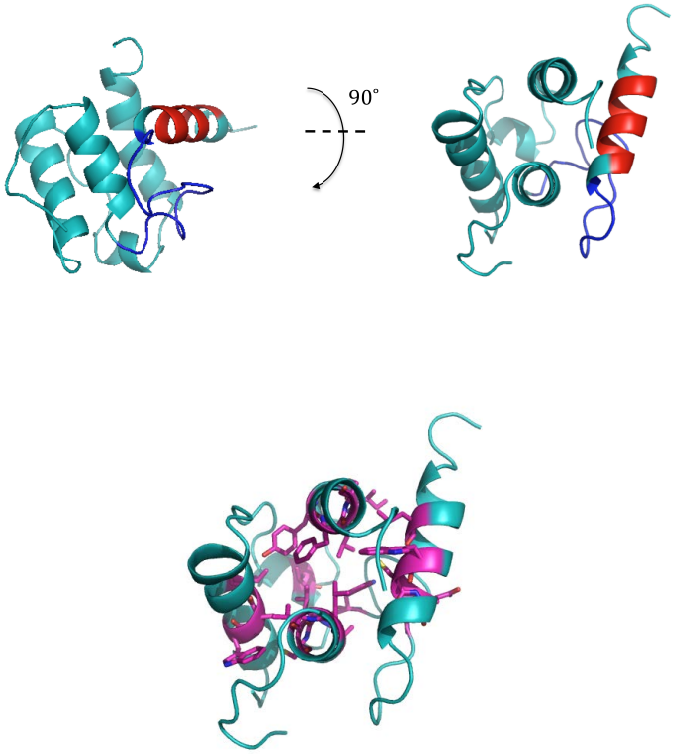
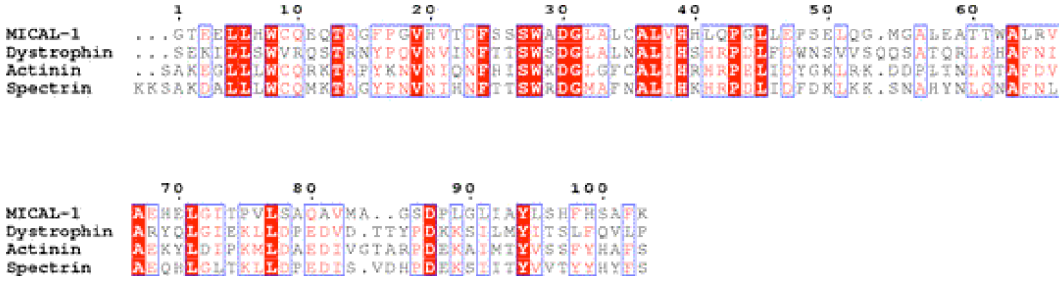


Figure 22. Overview of the CH domain from MICAL$_{FD-CH}$.

Upper panel: two views of the CH domain. Residues comprising the conserved actin-binding site found in type-2 CH domains are in red. Residues that match the PIP_2 binding site in type-2 CH domains are in blue. Lower panel: CH domain, highlighting the residues that make up the hydrophobic core – magenta, sticks.

A

B



C

Figure 23. Similarity of MICAL's CH domain to other type 2 CH domains.

(A) Structural alignments, all three align with an rms of ~ 1Å – CH domain of $MICAL_{FD\text{-}CH}$ (cyan) and type-2 CH domain from α-actinin (PDB 2EYI, pink), spectrin (PDB 1BKR, orange), and dystrophin (PDB 1DXX, blue). (B) Sequence alignment – CH domain of $MICAL_{FD\text{-}CH}$ with the same proteins in (A). (C) CH domain of $MICAL_{FD\text{-}CH}$ (cyan) and *human* MICAL-1 CH domain (PDB 1WYL).

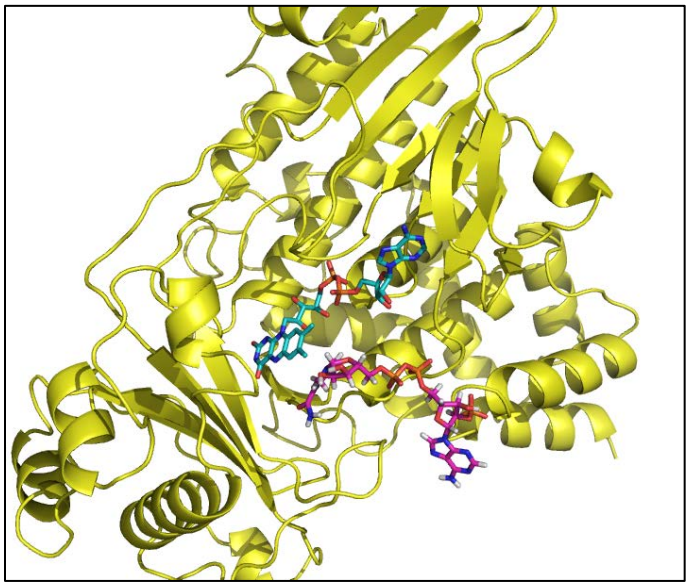


Figure 24. Model of NADPH bound to the FD of $MICAL_{FD\text{-}CH}$.

NADPH was docked by positioning the phosphate at the site where the electron density for Cl^- was observed in the $MICAL_{FD}$ crystal structure (PDB 2BRA).

4. Model of $MICAL_{FD\text{-}CH}$ Bound to Actin

The crystal structure of $MICAL_{FD\text{-}CH}$ (described in Chapter 3) reveals that the CH domain is connected to the FD with a disordered linker. This may be important for better accommodating the binding of the FD to actin, so that it can oxidize actin residue M44. Since it is technically challenging to obtain crystals of F-actin, or $MICAL_{FD\text{-}CH}$ bound to F-actin, this chapter describes the use of a computational approach to propose a model of $MICAL_{FD\text{-}CH}$ bound to F-actin.

4.1 Introduction

MICAL modifies F-actin by oxidation of a methionine (residue Met44) to methionine sulfoxide, and this is thought to destabilize the filament and lead to depolymerization [7]. It is unknown whether this oxidation is direct, where during the reaction Met44 is close enough to the C4a atom of the reduced FAD (the C4a-hydroperoxyflavin FAD intermediate) isoalloxazine ring (FAD-C4a), for direct transfer of the oxygen atom, or whether the oxidation is indirect, where the NADPH oxidase activity leads to a high local concentration of H_2O_2, and the H_2O_2 transfers an oxygen to the methionine [13, 14]. An atomic-resolution structure of $MICAL_{FD\text{-}CH}$ bound to F-actin would significantly aid in resolving this. F-actin is a polymer that can be very large; a solution of which will contain many polymers of different sizes. In addition to purity, proteins should be soluble and conformationally homogenous to yield crystals. F-actin has thus proven difficult to crystallize, and the highest resolution model obtained thus far is by x-ray fiber diffraction of F-actin polymers oriented in a magnetic field (PDB 2ZWH) [43]. Thus, it is technically challenging to obtain an atomic-resolution structure of a complex of $MICAL_{FD\text{-}CH}$ and F-actin.

The residue of actin that MICAL oxidizes, Met44, is well below the surface of F-actin. For a direct oxidation reaction to be possible, a conformation should be found in which the D-loop can reach to the active site of the FD in such a way that the Sulfur atom of the methionine is within ~3 Å of FAD-C4a. Direct oxidation cannot be possible if the Sulfur atom is too far away from FAD-C4a. This

conformation should be accessible from observed structures without breaking any covalent bonds or affecting the structural integrity of either actin or MICAL.

As described in the section 3.3, the actin D-loop is flexible and can adopt multiple conformations in solution, depending on the nucleotide state or the protein that it is bound to. Furthermore, in the crystal structures where it is visible, it is observed to adopt different conformations. The D-loop makes extensive electrostatic and hydrophobic contacts with the adjacent monomer in the F-actin filament [10, 41, 43], and chemical modifications in the D-loop dramatically decrease the thermal stability of F-actin [44]. Thus, even though the D-loop is flexible and able to adopt multiple conformations in monomeric actin as well as in the actin filament, its conformation and integrity is important for filament stability [41].

In this thesis, a computational approach was used to propose a model of this complex, and test whether direct oxidation of actin by MICAL is possible. With the crystal structure of $MICAL_{FD\text{-}CH}$ and a dimer from the highest resolution F-actin model (PDB 2ZWH), molecular dynamics simulations were performed. The simulations involved gentle steering of actin's M44 towards FAD-C4a, to assess whether the D-loop can adopt such a conformation.

4.2 Model Building

A model of the complex of $MICAL_{FD-CH}$ and F-actin was made using the crystal structure of $MICAL_{FD-CH}$ reported in this study and a dimer of actin from the most recent F-actin model (PDB 2ZWH), shown in Figure 25. The conformation of the FAD in the $MICAL_{FD-CH}$ structure was replaced by the one present in the crystal structure of the FD with the reduced FAD (PDB 2C4C). To obtain a starting model, the CH domain of $MICAL_{FD-CH}$ was aligned to the CH domain of an electron microscopy model of actinin bound to F-actin (PDB 3LUE), followed by manual adjustment of $MICAL_{FD-CH}$ to orient the opening of the active site (opposite the site where NADPH is predicted to bind, Figure 24) in the FD towards the D-loop of actin that contains M44, whilst maintaining the contact between the CH domain and actin. For direct oxidation to be possible, the D-loop must adopt a conformation in which the sulfur atom of M44 is within ~3 Å of FAD-C4a in MICALs active site.

4.2.1 Molecular Dynamics Simulation

To assess whether a conformation of the D-loop in which M44 is close to the FAD-C4a is possible, molecular dynamics calculations were performed (see chapter 5 for brief background on molecular dynamics). CHARMM version 28b2 with the CHARMM 28b2 [45] force field was used in the computations with implicit solvent and a distance dependent dielectric constant. For all calculations, the FAD cofactor was harmonically constrained to either the oxidized "out" conformation or the reduced "in" conformation with a force constant of 1 kcal/mol $Å^2$. The original

coordinates of this model were optimized by minimizing the energy for 1000 cycles of steepest descent, followed by 1000 cycles of conjugate gradient, and finally 1000 cycles of adopted-basis Newton-Raphson minimization. Leap frog Verlet molecular dynamics simulations were performed at constant temperature of 300 K and run for 30,000 fs, harmonically constraining M44 to a chosen position with a force constant of 0.5 kcal/mol $Å^2$. In each successive simulation, M44 was constrained to a position closer to the FAD-C4a atom, until the D-loop achieved a conformation where M44 was within the 3 Å range required for oxidation to be possible. The average value of the energy and its fluctuations during the last 10,000 fs were calculated with an in-house written program. Structural representations were prepared with the molecular visualization program Pymol [46].

4.3 Results

4.3.1 Model of the Arrangement for Direct Oxidation of Actin by MICAL

Upon reduction by NADPH, the FAD cofactor of MICAL undergoes a conformational change from the oxidized “out” conformation to the reduced “in” conformation that renders it less accessible to solvent. The molecular dynamics simulations reveal that it is possible to attain a D-loop conformation of actin in which M44 is in close proximity to the reduced FAD-C4a (hereafter referred to as “D-loop out” conformation), as shown in Figure 26. The major conformational change is in the D-loop itself; no other structural changes in other parts of the actin monomer are observed. The fact that this conformation can be reached in a molecular dynamics simulation, using gentle steering and without the need to break bonds or significantly alter the integrity of either structure, indicates that this conformation can be populated with a high enough frequency to be kinetically competent.

Given that the D-loop itself undergoes a large conformational change required to adopt the “D-loop out” conformation (Figure 26), the energetic cost of adopting this conformation was estimated by a series of short molecular dynamics simulations. The potential energy, averaged over the production phase of the simulation, was compared between simulations for an actin monomer where the D-

loop is in the “D-loop out” conformation (with FAD either oxidized or reduced) and an actin monomer whose D-loop is at the starting conformation found in the original PDB by Oda et al. The energy of the “D-loop out” conformation is within 2% of the original conformation (Table 3).

Table 3. Average energies from molecular dynamics simulations.

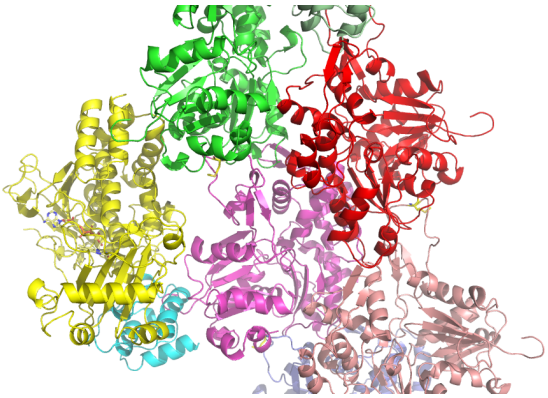
	Initial	D-loop out
Potential	3604.90	3548.62
Bond	1004.31	1009.02
Angle	2560.50	2552.89
Dihedral	1930.89	1918.92
Improper	177.15	178.00
Van der Waals	-1804.69	-1827.88
Electrostatic	-593.64	-609.05
Harmonic	3.97	2.42

All energies are reported in kilocalories per mole.

4.4 Conclusions and Discussion

Although our model suggests that direct oxidation of actin by MICAL is possible, it cannot be ruled out that actin Met44 oxidation occurs by a high local increase in H_2O_2 concentration upon binding. F-actin stimulates the NADPH oxidase activity of MICAL, and it is known that reduced FAD can either combine with O_2 to form a C4a hydroperoxidate that can release H_2O_2 or directly oxidize its substrate. This model does not rule out the possibility that Met44 is the residue of actin that is oxidized *directly*, and Met47 can become oxidized *indirectly* by H_2O_2. Both Met44 and Met47 are oxidized upon addition of MICAL and NADPH to F-actin; however, MICAL failed to depolymerize actin M44L mutants, but could still depolymerize actin M47L mutants [7]. Both these mutants were able to polymerize normally into F-actin and able to bind MICAL. Given the position of Met47 in the "D-loop out" conformation, this hypothesis is indeed possible. Future work using mutagenesis and electron microscopy studies would be valuable in addressing this possibility.

A



B

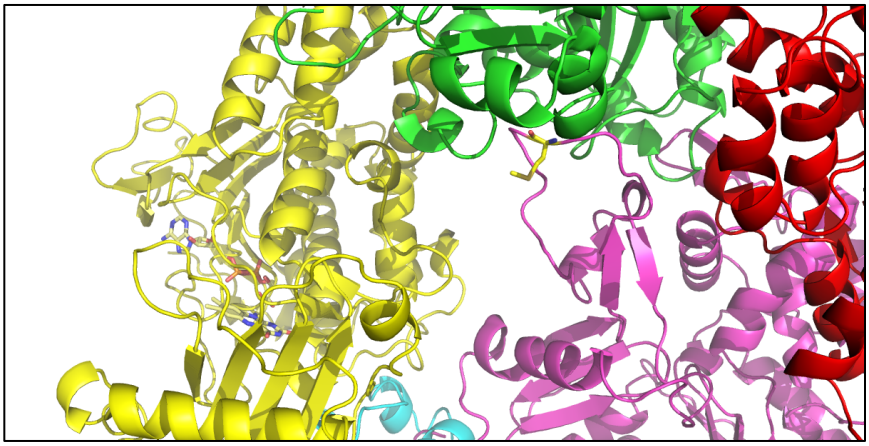


Figure 25. Model for MICAL$_{FD-CH}$ complex with F-actin.

(A) Overview of MICAL$_{FD-CH}$ interaction with F-actin in the model used for molecular dynamics simulations. FD is in yellow, CH in cyan, individual actin monomers from a polymer of F-actin (PDB 2ZWH) are in magenta, pink, red, and green. Actin's Met44 of the monomer in magenta is shown in yellow sticks. (B) Close-up of the figure in (A).

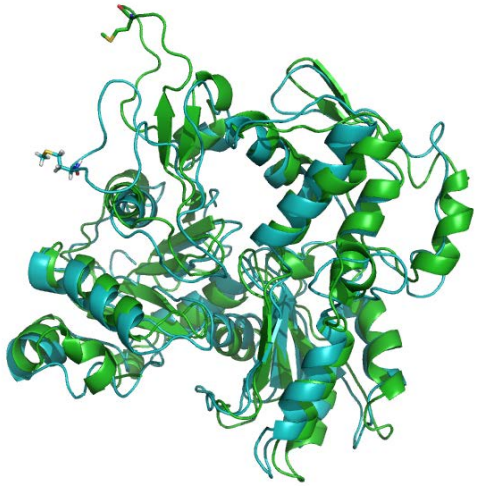


Figure 26. Possible model of conformational change in D-loop of actin.

This model shows a possible conformational change of the D-loop of actin upon binding $MICAL_{FD\text{-}CH}$, based on molecular dynamics simulations. The figure shows structural alignment of "D-loop out" conformation of the actin monomer whose D-loop changes conformation in the molecular dynamics based model (cyan) with respect to the initial conformation (green), highlighting that the rest of the monomer maintained its structural integrity.

5. Methods

In this thesis, structural characterization of $MICAL_{FD\text{-}CH}$ fragment was used to evaluate the effect of the presence of the CH domain on the FD. The most commonly used method for high-resolution structure determination of macromolecules is x-ray crystallography, and this was the method used to determine the structure of $MICAL_{FD\text{-}CH}$. Structure determination by x-ray crystallography is no trivial undertaking; there can be no structure without a protein crystal. In this chapter, I will describe at a very basic level the theory behind x-ray crystallography and its use in structure determination. In the final section I will discuss, at a very basic level, the theory behind computational molecular dynamics (MD) simulations. MD simulations were used in this thesis work to propose a model of $MICAL_{FD\text{-}CH}$ bound to F-actin.

5.1 X-ray Crystallography

Before introducing the method of x-ray crystallography, I will provide a brief account of “x-rays” and “crystals”, both of which are fundamental for ability to use this method. X-rays are a form of electromagnetic radiation of high energy, with wavelengths of 0.2 – 0.01 nm. Upon interaction with macromolecules, the electrons in the atoms of the macromolecule cause the x-rays to scatter from the original incident beam direction. This pattern of “diffracted” x-rays can be detected and measured to reconstruct the image of the scattering macromolecule [47] [48]. The diffraction pattern from a single macromolecule is too weak to detect, however. To circumvent this problem, crystals of a macromolecule are required. Crystals are a solid-state form of matter containing a large number of the same macromolecule, arranged in a periodic and ordered manner.

X-ray crystallography involves passing an x-ray beam through a crystal of a macromolecule. Since a crystal contains many macromolecules, each of which contains many atoms and their respective electrons, there are many scattered waves of x-rays. These waves either constructively or destructively interfere with each other, and the measured diffraction pattern from a diffraction dataset shows discrete points (with a given intensity) that are a result of constructive interference of the aforementioned waves (shifted with respect to each other by a certain phase).

A Fourier transform of the diffraction data is used to reconstruct the electron density of the scattering macromolecule. However, only the intensity of the discrete points can be obtained experimentally from the diffraction data; phases cannot be

obtained. This is the so-called "phase problem" in x-ray crystallography. This is addressed experimentally by addition of heavy atoms to use isomorphous replacement, or addition of anomalous scatterers to use anomalous diffraction. The phase problem can also be addressed without the need for additional experiments; if a model with greater than 30% identity and similar structure to the macromolecule from the crystal is available, the information from this search model can be used. This approach is called molecular replacement, and was successfully applied in this thesis work to obtain missing phases for structure determination. In this section, I will introduce very briefly some background on x-ray diffraction, the method of molecular replacement for phase determination, and refinement of the atomic models. For a more detailed discussion on the theory behind x-ray crystallography, I refer you to the textbook by Bernhard Rupp [49].

5.1.1 X-ray diffraction theory

Diffraction from a crystal

The complex structure factor $\mathbf{F_h}$, which describes the scattering in a given direction **h** by the summation of the scattering contributions f_j of j atoms at position $\boldsymbol{x_j}$ in the unit cell, can be expressed as:

$$F_h = \sum_{j}^{atoms} f_j \cdot \exp\left(2\pi i \mathbf{h}\mathbf{x}_j\right) \tag{5.1}$$

Bragg's Law

In order for reflections of the scattered x-rays to appear during data collection, they must be scattered at certain angles relative to the incident beam. If the scattering angle is an integer multiple of the wavelength, the resulting diffraction will be non-zero:

$$n\lambda = 2d_{hkl}\sin\theta \tag{5.2}$$

This is known as Bragg's law, and is a consequence of the periodicity of the macromolecules in the crystal.

Reconstructing the electron density

Rewriting the structure factor equation (5.x) in terms of spatial distribution over the entire volume V of the unit cell:

$$\mathbf{F}(hkl) = V\int_{x=0}^{1}\int_{y=0}^{1}\int_{z=0}^{1}\rho(xyz)\exp\left[2\pi i\left(hx+ky+lz\right)\right] \tag{5.3}$$

where $\rho(xyz)$ is the electron density at the coordinates xyz. Expressing $\rho(xyz)$ as a function of $\mathbf{F}(hkl)$, we obtain:

$$\rho(xyz) = 1/V\int_{h}\int_{k}\int_{l}|\mathbf{F}(hkl)|\exp\left[-2\pi i\left(hx+ky+lz\right)+i\alpha(hkl)\right] \tag{5.4}$$

where $\alpha(hkl)$ is the phase of the reflection (hkl). From the experimental data collection of diffraction from a crystal, only the intensity (discrete spots) can be obtained $I(hkl)\ \alpha\ |\mathbf{F}(hkl)|^2$. The phases need to be determined by one of the methods mentioned in the next section.

5.1.2 Phase Determination

In order to obtain the phases required for reconstructing the electron density of the scattering macromolecule, the most common methods are: (1) isomorphous replacement, (2) anomalous diffraction, and in the case where a structurally similar model is available (3) molecular replacement. Since in this thesis work the molecular replacement method was used, a very brief background is provided below.

5.1.2.1 Molecular replacement

Since this method does not require any experimental setup, and with the large number of structures currently available in the protein data bank (PDB), molecular replacement is the most commonly used method to obtain missing phases in structure determination. The condition is that there must be an existing structure with greater than 30% identity to use as a search model.

To determine the phases of an unknown structure, the search model with known phases is relocated and structure factors are calculated. This is repeated until the calculated structure factors of the search model are in agreement with the observed structure factors of the unknown structure. Although this six-dimensional search problem is fine in theory, macromolecules contain many atoms (on the order of 10,000) and searching each of the six degrees of freedom is computationally inefficient and expensive. A more efficient way to approach this problem is to use a

three-dimensional rotation search followed by a three-dimensional translation search.

In this way, the search begins with finding the proper orientation of the known structure in the crystal using a rotation function. Once this is achieved, the search for the correct position of the molecule in the unit cell is carried out using a translation function. After each step of the rotation and translation search, structure factors for the search model are calculated and evaluated against the structure factors of the unknown structure, until an acceptable agreement(s) is found.

Molecular replacement can only be used if a structure of a highly homologous model in a similar conformation is available. Success of this method also depends on the quality of the diffraction data obtained from the data collection.

5.1.3 Refinement

Once the phases are obtained, the electron density can be reconstructed, and model building ensues. To ensure that the model built is free of errors, the model undergoes several rounds of refinement. The agreement between the built model and the observed structure factors is determined using the R-value [50]:

$$R = \frac{\sum_h |F_{obs} - F_{calc}|}{\sum_h F_{obs}} \tag{5.5}$$

The R-value measures the difference between the observed structure factors from the diffraction data collected and the calculated structure factors from the built model. Refinement is carried out until convergence is achieved. To ensure that the

model does not contain any bias, the R-value is calculated in the same way as mentioned above, but with 5-10% of the observed reflections removed; this R-value is termed R-free. In an unbiased refinement, the difference between the R-value and R-free is minimal.

5.2 Molecular Dynamics Simulation

The atomic resolution structures of macromolecules determined by x-ray crystallography are extremely valuable in providing detailed structural insight into their function that cannot be obtained from experimental work alone. However, the structure represents a single conformation of a dynamic macromolecule. One of the most widely used methods to obtain dynamic information about macromolecules is Molecular Dynamics (MD) simulations. MD simulations are used to simulate motions of macromolecules as a function of time *in silico*. Information concerning structural and dynamic properties can be obtained from MD, such as molecular geometries, local fluctuations, and nature of various types of motions. This technique was used in this thesis work to determine whether a loop in a protein can move in a certain direction and adopt a certain conformation without the need for bond rearrangement or affecting the structural integrity of the protein [51]. I will present a very brief introduction to the theory underlying MD simulations.

Performing an MD simulation involves starting with an initial atomic model (from crystallographic or homology-modeling data), calculating the molecular forces acting upon each atom, moving each atom according to those forces, and finally advancing the simulation time by 1 fs. This process is repeated many times, and the atom positions and velocities at each timestep are stored for future analysis.

To simulate atomic motion, the equation that describes Newton's second law is integrated for each atom:

$$a_i(t) = \frac{d^2 x_i}{dt^2} = \frac{F(x_i)}{m_i} \tag{5.6}$$

where a_i is the acceleration of atom i at time t, determined by the force $F(x_i)$ acting on atom i of mass m_i at position x_i. The force $F(x_i)$ acting on each atom can be estimated, in the simplest case, by taking into account the contributions arising from bonded and non-bonded atoms [52, 53]:

$$E_{total} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}} \right] \tag{5.7}$$

Macromolecules are modeled as beads connected by simple springs; the beads represent the atoms, and the springs represent the covalent bonds that connect them. Dihedral angles are modeled using a sinusoidal function that approximates the energy differences between eclipsed and staggered conformations. Forces arising from non-bonded atoms as a result of Van der Waals interactions are modeled using the Lenard-Jones potential, and forces arising as a result of electrostatic interactions are modeled using Coulomb's law. To make the simulation a more accurate representation of the actual motion of real molecules, the energy terms above are parameterized to fit experimental data e.g. experimentally determined bond lengths and angles for carbon-carbon or carbon-nitrogen atoms. These parameters are what define the 'force field'; contributions to the force experienced by each atom in the field of the simulation.

In practice, a few steps are required to prepare a model for MD simulation. First, the initial atomic model must be placed in an explicit solvent environment, and following addition of missing atoms and assignment of charges and atom types, the model is minimized in order to optimize the coordinates towards a local energy minimum. Second, the system is slowly heated to reach the temperature required for the simulation. Third and finally before starting the simulation, the system is equilibrated to allow the system to relax from any high-energy conformation that may have come from the initial model.

A. Specificity of CasL and MICAL Interaction

MICAL was initially identified in a far-Western screening for proteins that interact with the SH3 domain of CasL (Crk-associated-substrate Lymphocyte type) [11]. Since then, no studies on the interaction of CasL and MICAL and its importance have been reported. CasL is a large multi-domain cytosolic protein that acts as a scaffold and is important for intracellular signaling in a variety of processes such as cell migration and cancer metastasis. In this thesis, the specificity of the interaction of CasL and MICAL was studied by Isothermal Titration Calorimetry (ITC), and a computational modeling approach was used to propose the reason behind the specificity.

A.1 Introduction

CasL is a cytosolic protein part of the p130Cas protein family (contains p130Cas, CasL, and EFS/SIN proteins), and acts as a scaffold that is heavily serine/tyrosine phosphorylated and important for normal and pathological cellular processes. CasL is important for transducing integrin-mediated signals, and has also been reported to function downstream of receptor tyrosine kinases, GPCRs, and chemokine receptors in T-cells and B-cells. Proper function of CasL is crucial for processes such as cell migration, cell adhesion, apoptosis, neurological development, proliferation and differentiation of normal and various types of cancer cells. Cas-L is highly conserved in all vertebrates; and no homologue exists in other lower organisms except *Drosophilla* (single Cas protein) [54-57].

CasL is 834 amino acids long, and contains multiple protein-protein interaction sites (Figure 27). It has an SH3 (Src-homology-3) domain at its N-terminus (aa 10-65), which can interact with other proteins that contain a poly-proline motif. The next domain downstream (aa ~90-350) is known as the "substrate domain" and contains repeats of tyrosine motifs that once phosphorylated act as binding sites for SH2 domain-containing proteins. Some studies have shown that p130Cas (CasL homologue with similar domain architecture and high sequence similarity) can act as a mechanosensor; extension of the substrate domain by an external force is sufficient to open up the tyrosine motifs and allow their phosphorylation, and thus the substrate domain itself is unstructured [58, 59]. The following region between amino acids 350-650 is a

serine-rich region that has a four-helix bundle structure, and is also an interaction site for other protein partners [60]. Finally, the carboxy-terminal domain that is highly conserved amongst all Cas family proteins has a FAT (Focal Adhesion Targeting) domain fold that also gets phosphorylated and forms binding sites for other partner proteins [61].

Integrin receptors are transmembrane proteins that mediate cell-cell and cell-matrix contacts at focal adhesions. Signalling through integrin receptors leads to signal transduction required for cell adhesion, migration, apoptosis, proliferation, and differentiation. Upon stimulation of integrins, CasL and/or p130Cas move from the cytosol to the focal adhesion region, where they perform a scaffolding function, physically linking many proteins together to transduce the required downstream signals and becoming hyper-phosphorylated in the process. CasL phosphorylation is in part influenced by the state of actin filaments, as disruption of actin filaments was shown to lead to dephosphorylation of CasL [55].

Cell migration and axon extension share parallel features such as coordination of actin dynamics in response to external signals. Studies show that CasL is phosphorylated in cerebral cortex and hippocampus neurons. Of particular relevance to MICAL and axon guidance, *Drosophilla* Cas (DCas) has been shown to be required for axon guidance during development [62]. Specifically, it functions downstream of integrin receptors and regulates the degree of fasciculation among axons. This function is highly dependent on the SH3 domain, and $DCas_{SH3}$ shares 62% sequence identity to $CasL_{SH3}$ domain. Interestingly, both *gain-of-function* and *loss-of-function* mutants of DCas showed a decrease in axon defasiculation compared

with the wild type, indicating that the function of DCas is *necessary* but not *sufficient* for proper axon guidance. This leads to the hypothesis that other factors or signaling cascades are required for proper axon defasiculation in addition to Integrin/DCas action. The authors of this study mention that their preliminary data suggests the physical interaction of MICAL and CasL is functionally important for axon guidance. To date, no studies on the interaction between MICAL and CasL have been reported.

In this thesis work, the initial aim was to characterize the effect of CasL on MICAL's redox activity. The work presented here addresses the specificity of the CasL-MICAL interaction. To our knowledge, the work outlined in this chapter represents the first biophysical characterization of this interaction. The interaction was studied by ITC and thermodynamic parameters of the binding were determined. Furthermore, the homology model of $CasL_{SH3}$ bound to a peptide of MICAL containing the poly-proline sequence reveals insight into the specificity of the interaction.

A.2 Methods

A.2.1 Expression and Purification of $CasL_{SH3}$

Cloning

A plasmid containing DNA coding for CasL from *mouse* (codon-optimized for expression in *Escherichia coli* – Genescript Inc.) residues 2-74 was cloned onto a pGEX-6P2 expression vector containing an N-terminal GST tag with an engineered N-terminal TEV protease site, and used to transform *Escherichia coli* BL21.

Protein expression and purification

Cells were grown for 4 hours at 37°C in LB media and protein expression was induced by addition of 0.5 mM IPTG after 3 hours, before harvesting by centrifugation. Cell pellets were lysed in 1X PBS buffer pH 7.3 containing 1mM DTT, 0.5 mM EDTA, and 2 complete EDTA-free protease inhibitor cocktail pills (Roche), and the protein was purified by GST-Sepharose affinity chromatography with a gradient of 50 mM Tris-HCl pH 8.0, 0-20 mM glutathione. The eluted protein was digested overnight with TEV protease to remove the GST-tag, and buffer exchanged to 50 mM Tris-HCl pH 7.0, 200 mM Glycine using a HiTrap Desalting column (GE Healthcare). The cleaved product was purified by using a Source 15Q anion exchange column, eluting with 50 mM Tris-HCl pH 7.0, and 0-1 M NaCl gradient (Figure 28).

Although a high yield (e.g. ~ 30 ml at ~ 3 mg/ml) of purified GST-tagged $CasL_{SH3}$ and/or untagged $CasL_{SH3}$ can be readily obtained, it was not possible to concentrate a sample of $CasL_{SH3}$ beyond 5 mg/ml, perhaps as a result of precipitation of the protein at the membrane of the typically used centrifugal filters. Attempts to concentrate the protein in a dialysis bag over PEG flakes also proved unfruitful. Lyophillization of the sample was also attempted, but the reconstituted protein solution was only as high as ~ 8 mg/ml. This could be further optimized, testing with an appropriate volatile buffer prior to lyophilizing.

A.2.2 Expression and Purification of CasL

Cloning

A plasmid containing DNA coding for CasL from *mouse* (codon-optimized for expression in *Escherichia coli* – Genescript Inc.) was cloned onto a pGEX-6P2 expression vector containing an N-terminal GST tag with an engineered N-terminal TEV protease site, and used to transform *Escherichia coli* BL21.

Protein expression and purification

Cells were grown for 4 hours at 37°C in LB media and protein expression was induced by addition of 0.5 mM IPTG after 3 hours, before harvesting by centrifugation. Cell pellets were lysed in 1X PBS buffer pH 7.3 containing 0.1 % Tween-20, 2 mM DTT, 5 mM $MgCl_2$, 10 % v/v glycerol, and 2 complete EDTA-free protease inhibitor cocktail pills (Roche), and the protein was purified by GST-

Sepharose affinity chromatography with a gradient of 50 mM Tris-HCl pH 8.0, 0-20 mM glutathione. This step usually yields ~ 45 ml of ~ 0.6 mg/ml of GST-tagged CasL. The eluted protein was digested overnight with TEV protease to remove the GST-tag, and buffer exchanged to 1X PBS buffer pH 7.3 (Figure 29). The cleaved product was purified first by GST-sepharose affinity chromatography to remove uncut protein and the GST tag, followed by buffer exchange to 50 mM Tris-HCl pH 7.0, 200 mM sucrose, 10 % v/v glycerol and finally by using a Source 15Q anion exchange column, eluting with 50 mM Tris-HCl pH 7.0, and 0-1 M NaCl gradient (Figure 30). This step usually yields ~ 30 ml of ~ 0.5 mg/ml.

A.2.3 Isothermal Titration Calorimetry

ITC is a method where both a sample cell and reference cell are kept at constant slow temperature jump, and upon injection of a protein or small molecule (e.g. ligand, peptide, drug, etc.) that binds with a macromolecule in the sample cell, the heat required to keep the temperature of the sample and reference cells the same is recorded. A number of successive injections of injectant are performed. From this experiment, thermodynamic parameters such as equilibrium binding constant, free energy, enthalpy, and entropy of binding can be determined.

ITC experiments were carried out using 120 μM GST-tagged $CasL_{SH3}$ and a peptide containing either wild type MICAL's poly-proline sequence (SGAEPPPKPPRS) or peptides containing mutated sequence where an Ala, Glu, or Trp replaces the Lys. A VP-ITC Microcal calorimeter was used, and the experiment

was carried out at 27˚C. All reagents were in 50 mM HEPES pH 7.3, 200 mM NaCl. The sample cell contained 2 ml of the GST-tagged $CasL_{SH3}$, and 1.2 mM of MICAL peptide (wild-type or mutant) was added over 25 injections of 10 µl each.

A.2.4 Homology Modeling of CasL and MICAL

A homology model of $CasL_{SH3}$ was built based on the crystal structure of the SH3 domain of c-Crk bound to a peptide consisting of a poly-proline sequence from its protein interactor (PDB 1CKB) [63]. The model was built using the software MOE. A peptide consisting of the poly-proline sequence of MICAL PPPKPPR was modeled and minimized in MOE, and aligned to the poly-proline peptide originally found in 1CKB. Figures were made using the molecular visualization program Pymol [46].

A.3 Results

A.3.1 Binding of CasL to MICAL

The binding of CasL and MICAL was studied by ITC, using the SH3 domain of CasL as the macromolecule, and a peptide containing MICAL's poly-proline sequence as the injectant. Binding occurs with a K_D of ~ 175 μM, a favorable enthalpy of binding – 8 kcal mol^{-1}, and an unfavorable entropy upon binding of – 9.5 cal K^{-1} mol^{-1}. The unfavorable entropy is most likely due to the sequence being on a peptide that has a large conformational entropy, and not part of the full-length MICAL protein where it will have significantly less conformational entropy (Figure 31).

When peptides containing mutant MICAL poly-proline sequences (Lys to Ala, Glu, or Trp) were tested for binding to $CasL_{SH3}$ by ITC, all mutant peptides either failed to bind, or bound with very low affinity that is not detectable by ITC (Figure 31). This implicates the Lysine residue in the poly-proline sequence as key for the specificity of the interaction. Binding to the mutant peptides is exothermic as opposed to endothermic for the wild-type peptide, but the enthalpy is on the order of 0.1 μcal corresponding to the heat of dilution of the peptide.

A.3.2 Homology Model: Residues Potentially Important for Recognition

A homology model of $CasL_{SH3}$ bound to MICAL's poly-proline sequence peptide was constructed using the crystal structure of the SH3 domain of c-Crk bound to its poly-proline peptide as a template (PDB 1CKB). The model shows structural features of the binding site that could explain the specificity of the interaction and the importance of the Lys in MICAL's poly-proline sequence in this interaction.

Structural analysis of c-Crk_{SH3} bound to its PPVPP peptide shows the hydrophobic nature of the interaction (Figure 32). The Pro, Tyr, and Phe residues along with the 6-membered-ring of the Trp side chain provide a hydrophobic environment to accommodate the Val from the poly-proline peptide. This SH3 domain would not be able to accommodate a Lys in the position of the Val because of the hydrophobic environment and steric clashes as the Lys has a larger side chain.

Analysis of the model of $CasL_{SH3}$ shows that there is a deeper binding site ideal for a Lys side chain, and a less hydrophobic/more polar environment, evident upon alignment with c-Crk_{SH3}. Specifically, the Tyr of c-Crk_{SH3} is an Arg instead in $CasL_{SH3}$; the bulky Phe side chain of c-Crk_{SH3} is an Asn instead in $CasL_{SH3}$, and the Trp side chain is shifted outwards making the binding pocket larger. Furthermore, the side chain of an Asp residue inside the binding pocket of $CasL_{SH3}$ can adopt a conformation pointing towards the Lys of the poly-proline peptide, forming a potential salt bridge.

A.4 Conclusions and Discussion

Both an experimental and a computational approach were used in order to gain insight into the thermodynamics and specificity of the interaction between MICAL and CasL.

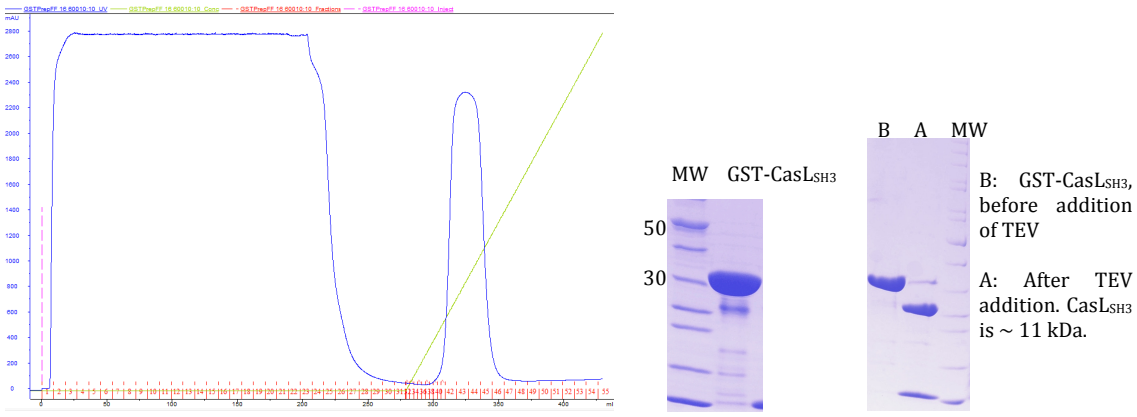
Experimentally, the binding of a peptide of MICAL's poly-proline sequence to $CasL_{SH3}$ was monitored by ITC. Mutational analysis implicates the Lys in the PPKPP sequence of MICAL as critical for the interaction, since the single substitution of the Lys for Ala (smaller side chain), Glu (opposite charge of Lys), or Trp (larger, bulkier side chain) was enough to abolish binding. This suggests that the binding pocket of CasL's SH3 domain is less hydrophobic and larger than that of c-Crk. The poly-proline motifs that typically bind SH3 domains are hydrophobic, with a general consensus sequence PXXP, and hence hydrophobic forces drive the interaction of SH3 domains with their respective poly-proline sequence in the target interacting protein. For example, c-Crk_{SH3} domain binds to poly-proline sequence (PPVPP) whereas $CasL_{SH3}$ binds to MICAL's poly-proline sequence (PPKPP) containing a Lys instead of the Val.

To gain further support for the SH3 domain of CasL having a different binding pocket than other hydrophobic SH3 domains, a computational approach was used to build a homology model of $CasL_{SH3}$ and bound to a peptide of MICAL's poly-proline sequence. The model reveals that the binding pocket of $CasL_{SH3}$ is indeed larger and more polar than that of c-Crk_{SH3}, and this could be the reason behind the specificity of the interaction between CasL and MICAL.

Figure 27. Overview of CasL domains.

The N-terminus consists of an SH3 domain, followed by an unstructured substrate domain that becomes heavily tyrosine phosphorylated during signal transduction. Next lies a Ser rich region that forms a four-helix bundle fold. The C-terminus has a FAT-like domain fold and interacts with many proteins including c-Crk.

A



B

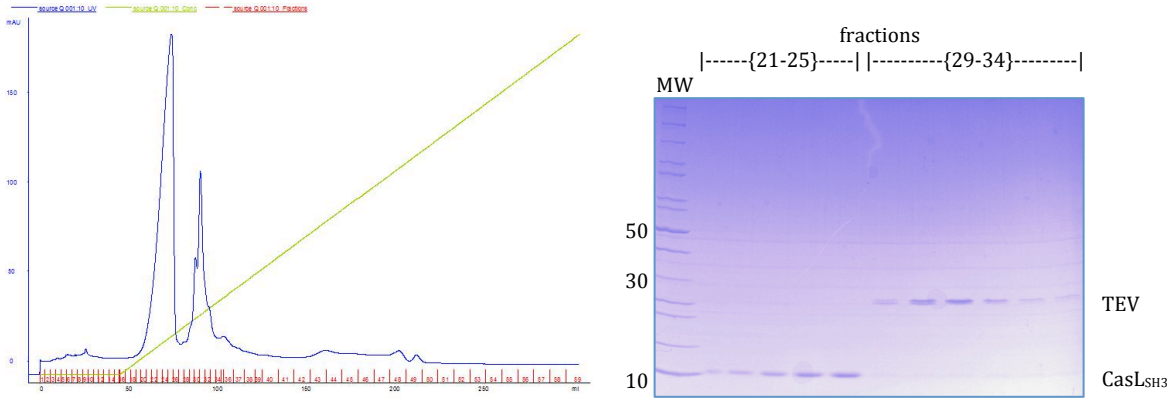
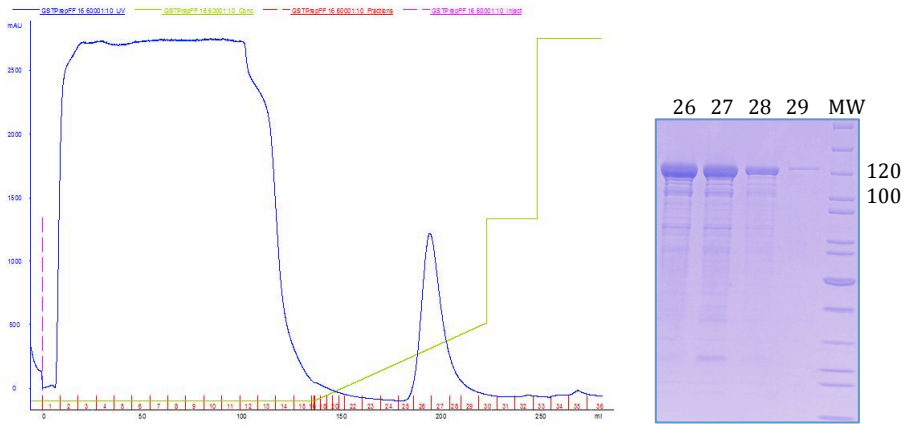


Figure 28. Purification of CasL$_{SH3}$.

(A) Purification of CasL$_{SH3}$ with glutathione sepharose affinity chromatography. Left: chromatogram from a GSTPrep™ FF 16/10 column (GE). Middle: SDS-PAGE analysis of the purification from this step. MW: Molecular weight marker in kDa. GST-CasL$_{SH3}$ represents a mixture of all the fractions corresponding to the peak. Right: SDS-PAGE analysis of the GST tag removal using TEV protease. (B) Left: chromatogram from a SOURCE15Q purification. Right: SDS-PAGE analysis of the purification from this step.

A



B

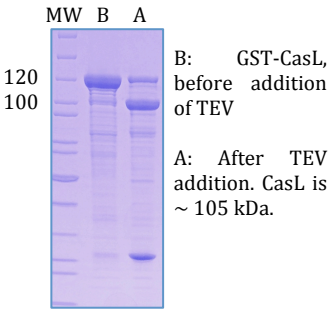
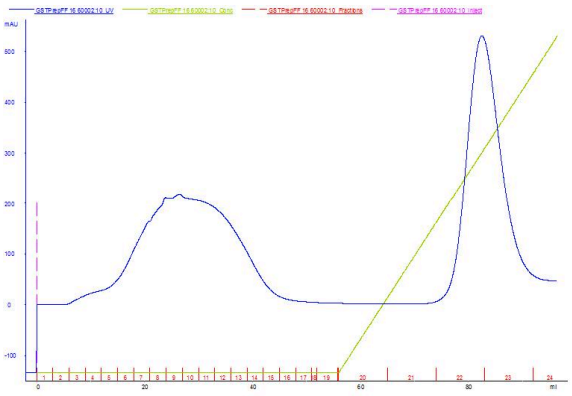
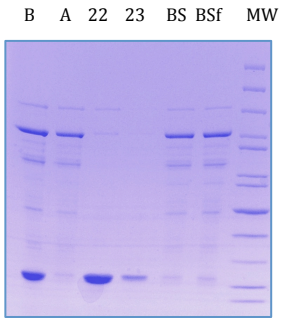


Figure 29. Purification of CasL: Affinity chromatography and tag cleavage.

(A) Purification of CasL with glutathione sepharose affinity chromatography. Left: chromatogram from a GSTPrep™ FF 16/10 column (GE). right: SDS-PAGE analysis of the purification from this step. MW: Molecular weight marker in kDa. (B) SDS-PAGE analysis of the GST tag removal using TEV protease.

A





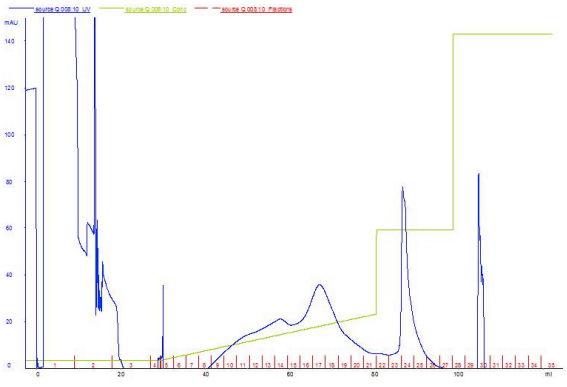
B: Sample loaded into column

A: fractions 6-14, containing CasL only

BS: Before SOURCE15S

BSf: filtered

B



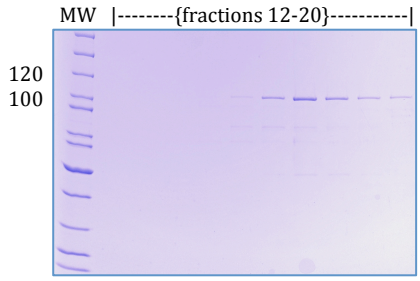
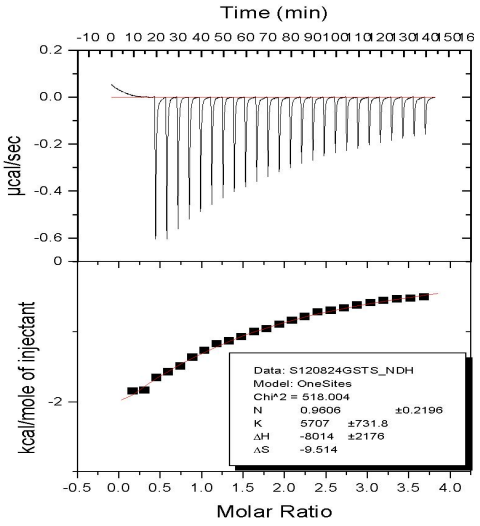


Figure 30. Purification of CasL: Ion exchange after tag removal.

(A) Purification of cleaved CasL, TEV, GST tag, and GST-CasL with glutathione sepharose affinity chromatography. Left: chromatogram from a GSTPrep™ FF 16/10 column (GE). right: SDS-PAGE analysis of the purification from this step. (B) Left: chromatogram from a SOURCE15Q purification. Right: SDS-PAGE analysis of the purification from this step.

A



B

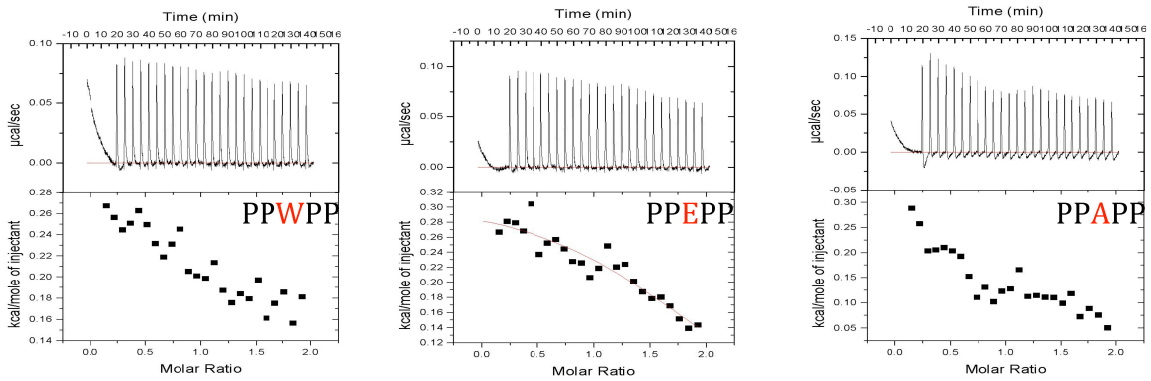


Figure 31. Isothermal titration calorimetry of CasL$_{SH3}$ with MICAL peptide.

(A) Figure shows the heat change of binding upon each injection of MICAL peptide. (B) Same as A, but peptide contains single amino acid substitution of the Lys (PPKPP) to Trp, Glu, or Ala (indicated in each curve inset).

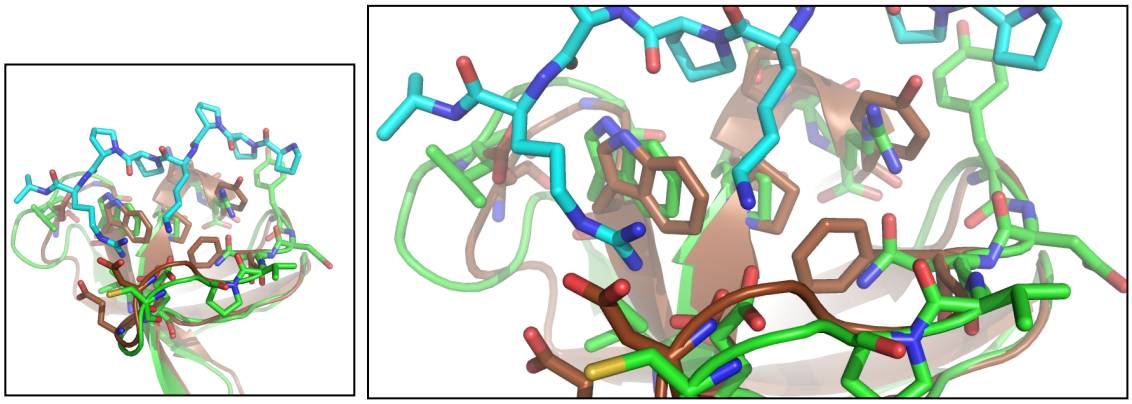


Figure 32. Homology model of $CasL_{SH3}$.

The right panel is a close-up of the left panel. The SH3 domain of c-Crk is shown in brown (PDB 1CKB), while $CasL_{SH3}$ model is shown in green. The PPKPP peptide (MICAL's poly-proline sequence) is shown in cyan. The sticks shown in the SH3 domains represent the residues in the binding pocket.

Bibliography

1. Tessier-Lavigne M & Goodman CS (1996) The molecular biology of axon guidance. *Science* 274(5290):1123-1133.
2. Dickson BJ (2002) Molecular mechanisms of axon guidance. *Science* 298(5600):1959-1964.
3. Kolodkin AL & Tessier-Lavigne M (2011) Mechanisms and molecules of neuronal wiring: a primer. *Cold Spring Harb Perspect Biol* 3(6).
4. Geraldo S & Gordon-Weeks PR (2009) Cytoskeletal dynamics in growth-cone steering. *J Cell Sci* 122(Pt 20):3595-3604.
5. Terman JR, Mao T, Pasterkamp RJ, Yu HH, & Kolodkin AL (2002) MICALs, a family of conserved flavoprotein oxidoreductases, function in plexin-mediated axonal repulsion. *Cell* 109(7):887-900.
6. Hung RJ*, et al.* (2010) Mical links semaphorins to F-actin disassembly. *Nature* 463(7282):823-827.
7. Hung RJ, Pak CW, & Terman JR (2011) Direct redox regulation of F-actin assembly and disassembly by Mical. *Science* 334(6063):1710-1713.

8. Prokop A, Beaven R, Qu Y, & Sanchez-Soriano N (2013) Using fly genetics to dissect the cytoskeletal machinery of neurons during axonal growth and maintenance. *J Cell Sci* 126(Pt 11):2331-2341.

9. Kalil K & Dent EW (2005) Touch and go: guidance cues signal to the growth cone cytoskeleton. *Curr Opin Neurobiol* 15(5):521-526.

10. Holmes KC, Popp D, Gebhard W, & Kabsch W (1990) Atomic model of the actin filament. *Nature* 347(6288):44-49.

11. Suzuki T*, et al.* (2002) MICAL, a novel CasL interacting molecule, associates with vimentin. *J Biol Chem* 277(17):14933-14941.

12. Hung RJ & Terman JR (2011) Extracellular inhibitors, repellents, and semaphorin/plexin/MICAL-mediated actin filament disassembly. *Cytoskeleton (Hoboken)* 68(8):415-433.

13. Vanoni MA, Vitali T, & Zucchini D (2013) MICAL, the Flavoenzyme Participating in Cytoskeleton Dynamics. *Int J Mol Sci* 14(4):6920-6959.

14. Giridharan SS & Caplan S (2013) MICAL-Family Proteins: Complex Regulators of the Actin Cytoskeleton. *Antioxid Redox Signal*.

15. Zhou Y, Gunput RA, Adolfs Y, & Pasterkamp RJ (2011) MICALs in control of the cytoskeleton, exocytosis, and cell death. *Cell Mol Life Sci*.

16. Nadella M, Bianchet MA, Gabelli SB, Barrila J, & Amzel LM (2005) Structure and activity of the axon guidance protein MICAL. *Proc Natl Acad Sci U S A* 102(46):16830-16835.

17. Zucchini D, Caprini G, Pasterkamp RJ, Tedeschi G, & Vanoni MA (2011) Kinetic and spectroscopic characterization of the putative monooxygenase domain of human MICAL-1. *Arch Biochem Biophys* 515(1-2):1-13.

18. Gimona M, Djinovic-Carugo K, Kranewitter WJ, & Winder SJ (2002) Functional plasticity of CH domains. *FEBS Lett* 513(1):98-106.

19. Zheng Q & Zhao Y (2007) The diverse biofunctions of LIM domain proteins: determined by subcellular localization and protein-protein interaction. *Biol Cell* 99(9):489-502.

20. Siebold C, *et al.* (2005) High-resolution structure of the catalytic region of MICAL (molecule interacting with CasL), a multidomain flavoenzyme-signaling molecule. *Proc Natl Acad Sci U S A* 102(46):16836-16841.

21. Sun H, *et al.* (2006) Solution structure of calponin homology domain of Human MICAL-1. *J Biomol NMR* 36(4):295-300.

22. Jin X, *et al.* (2007) Investigation of the four cooperative unfolding units existing in the MICAL-1 CH domain. *Biophys Chem* 129(2-3):269-278.

23. McDonald CA, Liu YY, & Palfey BA (2013) Actin stimulates reduction of the MICAL-2 monooxygenase domain. *Biochemistry* 52(35):6076-6084.

24. Friedberg F (2010) Singlet CH domain containing human multidomain proteins: an inventory. *Mol Biol Rep* 37(3):1531-1539.

25. Schmidt EF, Shim SO, & Strittmatter SM (2008) Release of MICAL autoinhibition by semaphorin-plexin signaling promotes interaction with collapsin response mediator protein. *J Neurosci* 28(9):2287-2297.

26. Giridharan SS, Rohn JL, Naslavsky N, & Caplan S (2012) Differential regulation of actin microfilaments by human MICAL proteins. *J Cell Sci* 125(Pt 3):614-624.

27. Winterbourn CC (2013) The biological chemistry of hydrogen peroxide. *Methods Enzymol* 528:3-25.

28. Lundquist MR*, et al.* (2014) Redox Modification of Nuclear Actin by MICAL-2 Regulates SRF Signaling. *Cell* 156(3):563-576.

29. Schneider CA, Rasband WS, & Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671-675.

30. Pasterkamp RJ & Verhaagen J (2006) Semaphorins in axon regeneration: developmental guidance molecules gone wrong? *Philos Trans R Soc Lond B Biol Sci* 361(1473):1499-1511.

31. Tran TS, Kolodkin AL, & Bharadwaj R (2007) Semaphorin regulation of cellular morphology. *Annu Rev Cell Dev Biol* 23:263-292.

32. Phillips K & de la Pena AH (2011) The combined use of the Thermofluor assay and ThermoQ analytical software for the determination of protein stability and buffer optimization as an aid in protein crystallization. *Curr Protoc Mol Biol* Chapter 10:Unit10 28.

33. Anonymous (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50(Pt 5):760-763.

34. Winn MD*, et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235-242.

35. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486-501.

36. Pelikan M, Hura GL, & Hammel M (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering. *Gen Physiol Biophys* 28(2):174-189.

37. Eswar N*, et al.* (2006) Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics* Chapter 5:Unit 5 6.

38. Sjoblom B, Ylanne J, & Djinovic-Carugo K (2008) Novel structural insights into F-actin-binding and novel functions of calponin homology domains. *Curr Opin Struct Biol* 18(6):702-708.

39. Galkin VE, Orlova A, Salmazo A, Djinovic-Carugo K, & Egelman EH (2010) Opening of tandem calponin homology domains regulates their affinity for F-actin. *Nat Struct Mol Biol* 17(5):614-616.

40. Durer ZA*, et al.* (2012) Structural states and dynamics of the D-loop in actin. *Biophys J* 103(5):930-939.

41. Dominguez R & Holmes KC (2011) Actin structure and function. *Annu Rev Biophys* 40:169-186.

42. Pivovarova AV, Khaitlina SY, & Levitsky DI (2010) Specific cleavage of the DNase-I binding loop dramatically decreases the thermal stability of actin. *FEBS J* 277(18):3812-3822.

43. Oda T, Iwasa M, Aihara T, Maeda Y, & Narita A (2009) The nature of the globular- to fibrous-actin transition. *Nature* 457(7228):441-445.

44. Oztug Durer ZA, Diraviyam K, Sept D, Kudryashov DS, & Reisler E (2010) F-actin structure destabilization and DNase I binding loop: fluctuations mutational cross-linking and electron microscopy analysis of loop states and effects on F-actin. *J Mol Biol* 395(3):544-557.

45. Brooks BR*, et al.* (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* 30(10):1545-1614.

46. Schrodinger, LLC (2010) The PyMOL Molecular Graphics System, Version 1.3r1.).

47. Bragg WH & Bragg WL (1918) *X rays and crystal structure* (G. Bell, London,) 3d Ed pp vii, 228 p.

48. Bragg WL (1975) *The development of x-ray analysis* (G. Bell, London) pp viii, 270 p.

49. Rupp B (*Biomolecular crystallography : principles, practice, and application to structural biology* (Garland Science, New York) pp xxi, 809 p.

50. Brunger AT (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355(6359):472-475.

51. Schlick T (2006) *Molecular modeling and simulation : an interdisciplinary guide* (Springer, New York) [Repr. with corrections] Ed pp xliii, 634 p.

52. Durrant JD & McCammon JA (2011) Molecular dynamics simulations and drug discovery. *BMC Biol* 9:71.

53. Nurisso A, Daina A, & Walker RC (2012) A practical introduction to molecular dynamics simulations: applications to homology modeling. *Methods Mol Biol* 857:137-173.

54. Seo S, Ichikawa M, & Kurokawa M (2006) Structure and function of cas-L and integrin-mediated signaling. *Crit Rev Immunol* 26(5):391-406.

55. Singh M, Cowell L, Seo S, O'Neill G, & Golemis E (2007) Molecular basis for HEF1/NEDD9/Cas-L action as a multifunctional co-ordinator of invasion, apoptosis and cell cycle. *Cell Biochem Biophys* 48(1):54-72.

56. Barrett A, Pellet-Many C, Zachary IC, Evans IM, & Frankel P (2013) p130Cas: a key signalling node in health and disease. *Cell Signal* 25(4):766-777.

57. O'Neill GM, Seo S, Serebriiskii IG, Lessin SR, & Golemis EA (2007) A new central scaffold for metastasis: parsing HEF1/Cas-L/NEDD9. *Cancer Res* 67(19):8975-8979.

58. Sawada Y*, et al.* (2006) Force sensing by mechanical extension of the Src family kinase substrate p130Cas. *Cell* 127(5):1015-1026.

59. Lu C, Wu F, Qiu W, & Liu R (2013) P130Cas substrate domain is intrinsically disordered as characterized by single-molecule force measurements. *Biophys Chem* 180-181:37-43.

60. Briknarova K*, et al.* (2005) The serine-rich domain from Crk-associated substrate (p130cas) is a four-helix bundle. *J Biol Chem* 280(23):21908-21914.

61. Mace PD*, et al.* (2011) NSP-Cas protein structures reveal a promiscuous interaction module in cell signaling. *Nat Struct Mol Biol* 18(12):1381-1387.

62. Huang Z, Yazdani U, Thompson-Peer KL, Kolodkin AL, & Terman JR (2007) Crk-associated substrate (Cas) signaling protein functions with integrins to

specify axon guidance during development. *Development* 134(12):2337-2347.

63. Wu X*, et al.* (1995) Structural basis for the specific interaction of lysine-containing proline-rich peptides with the N-terminal SH3 domain of c-Crk. *Structure* 3(2):215-226.

Curriculum Vitae

Saif Al Qassim

Department of Biophysics & Biophysical Chemistry
Johns Hopkins University, School of Medicine
725 N. Wolfe St., Wood Basic Science Building 604
Baltimore, MD 21205

Education

Ph.D.	2014	BCMB (Biochemistry, Cell & Molecular Biology) Department of Biophysics & Biophysical Chemistry Johns Hopkins School of Medicine Mentors: Mario Amzel, PhD and Mario Bianchet, PhD
B.S.	2008	Cell & Molecular Biology University of Michigan, Ann Arbor

Professional Experience

Graduate Research Assistant Johns Hopkins School of Medicine	2009-2014	Laboratory of Dr. Mario Amzel
Research Rotation Johns Hopkins University	2009	Laboratory of Dr. Doug Barrick
Research Rotation Johns Hopkins University	2008	Laboratory of Dr. Jeffery Gray

Scholarships

Scholarship Coordination Office, Office of the President of the United Arab Emirates (UAE), Abu Dhabi, UAE February 2014
Travel award

Scholarship Coordination Office, Office of the President of the United Arab Emirates (UAE), Abu Dhabi, UAE August 2008 – May 2014
Full graduate scholarship including tuition and stipend
Johns Hopkins School of Medicine

Scholarship Coordination Office, Office of the President of the United Arab Emirates (UAE), Abu Dhabi, UAE August 2004 – May 2008
Full undergraduate scholarship including tuition and stipend
University of Michigan, Ann Arbor

Awards and Honors

Poster prize, BCMB Annual Retreat, BCMB program, Johns Hopkins School of Medicine, Baltimore, MD, October 2012.

Publications, peer reviewed

Alqassim S. S., Amzel L. M., Bianchet M. A. Modulation of MICAL Activity by the Calponin Homology Domain: Structural Insights. *Manuscript in preparation.*

Alqassim S. S., Urquiza M., Bianchet V. A., Borgnia E., Amzel L. M., Bianchet M. A. How Non-redox Domains of MICAL-1 Modulate its Enzymatic and Actin-binding Activities. *Manuscript in preparation.*

Simon O., **Alqassim S. S**., Urquiza M., Bianchet M. A., Amzel L. M.The Plexin-A1 Intracellular Domain and Rnd1 Together Suffice to Lift Autoinhibition of MICAL-1's Monooxygenase, but not its Actin-Disassembling Activities. *Manuscript in preparation*.

Posters and Abstracts

Alqassim S. S., Bianchet M. A., Amzel L. M. Structural Insights into Modulation of MICAL Activities by its Calponin Homology (CH) Domain. Biophysical Society 58th Annual Meeting. San Francisco, CA, February 16, 2014.

Alqassim S. S., Bianchet M. A., Amzel L. M. Modulation of MICAL Activity by the Calponin Homology (CH) Domain. Institute for Biophysical Research, Baltimore, MD, September, 2013.

Alqassim S. S., Simon O., Urquiza M., Bianchet M. A., Amzel L. M. Regulation of the Redox Activity of MICAL. BCMB Annual Retreat, Baltimore, MD, October, 2012.

Alqassim S. S., Simon O., Urquiza M., Bianchet M. A., Amzel L. M. MICAL Acting on Actin: Kinetic Studies. Institute for Biophysical Research, Baltimore, MD, September, 2012.

Alqassim S. S., Bianchet M. A., Amzel L. M. MICAL Acting on Actin. Institute for Biophysical Research, Baltimore, MD, September, 2011.

Oral presentations

"An Introduction to the Theory of Protein Crystal Nucleation." Department of Biophysics & Biophysical Chemistry, Johns Hopkins School of Medicine, Baltimore, MD, November 2012.

"Role of Calponin Homology (CH) Domain in Regulation of MICAL Activities." BCMB Annual Retreat, Johns Hopkins School of Medicine, Baltimore, MD, October 2011.

"Role of Calponin Homology (CH) Domain in Regulation of MICAL Activitiy in Axon Guidance." BCMB Colloquium, Johns Hopkins School of Medicine, Baltimore, MD, October 2011.

Service and Leadership

Laboratory Research Mentor, Eitan Borgnia, high school student. Laboratory of Dr. Mario Amzel, Johns Hopkins School of Medicine. September 2013 – January 2014.

Tutor, Biophysics and Biophysical Chemistry, Johns Hopkins School of Medicine. September 2013 – January 14.

Writer, BCMB Community Newsletter, BCMB program, Johns Hopkins School of Medicine, Baltimore, MD. http://bcmbnews.com/author/saifaq/. September 2012 – February 2013.

Tutor, Biophysics and Biophysical Chemistry, Johns Hopkins School of Medicine. October 2012.

Committee head, for selection and invitation of BCMB Annual Retreat keynote speaker, Johns Hopkins School of Medicine, Baltimore, MD. October 2012.

Student emcee, BCMB Annual Retreat, Johns Hopkins School of Medicine, Baltimore, MD. October 2012.

Committee member, for selection and invitation of BCMB Annual Retreat keynote speaker, Johns Hopkins School of Medicine, Baltimore, MD. October 2011.

Tutor, Organic Mechanisms in Biology, Johns Hopkins School of Medicine. January 2011.

Professional Affiliations

Biophysical Society 2011-Present